

```
/organism="Triticum aestivum"  
/cultivar="Florida"
```

CDS	/db.xref="taxon:4565"	
	join(<3..296,397..1617,2145..2924)	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAC07514.1"	
	/db.xref="gi:9997325"	
	/translation="SGPRRLRRRRPRNATAGKGVENCAAVEATKYEDGCEDEPEV AEDRIALAGACRVLAGAPFLGATATLALGAVFAYSGGATATLCLTPEDLAKDRVTT EEVLDPLMNRGTVMVHVFIEGELNNMLYGRFEGGTAPHCGLDLVSNVVDYARAK VISGEYXVPARGNNCMPOAMGMLPLRPSTDEMGDPLRYPQDLVYEHLHGFPR HDSNVHEHPGFICAVSLKLDYLKELVGVCIELMGCHERNELEYSTSSKMFPGYST NFPSPMTRTSGGICKNGRDAINFEKPRVRAHNGIEIYLDVYFNHAEENEGPILII SFKVDNTTYMLAPKGEFYNSSCGATFNCHNVYRFYDCLARYVTBHHVDFRR DLASILTRGSSLPMDPVNRYAPLEGDMTCTPLVTPPLIMISNDPLGGVKLIAEG WDAGLYOVQGFPRMNVSEWNGKRYDVRPFIKGTGDFAGFEGELCGSHLYOAGG KPMHSINFCVAHDGFTLADLVYTNKKNALPGLNNRNPENHNLSMNGEEDFARLS VKRIRKQMRNFEVCLWYASGVPMPEVMEDEYGHNGKNNNTYCHDSYVNYPRMKKEE YSELNRCCIMLTFRKKEGEGLEDPEAPKLOMHHQPGKPDMSENSRFYAPSMKDE RQELTYAENTSHPAYVELPERAGRRMEPVYDGRPAFYDFLTDDLPRALITHQPS HFLSNLYPMLSYSSVILVLRPDV"	
BASE COUNT	803 a	636 c 762 g 796 t
ORIGIN		
Query Match	100.0%;	Score 2997; DB 6; Length 2997;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 2997; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1	ggtcgggacggcgccgcgcgtcgcagcgttgccgaccccaatgcgcgcggcggaaggcg
Db	1	ggtccggcgccggcgccgcgcgtcgcagcggctggcgaaccCAATCGAGCGGGGAAAGGGG
OY	61	tccggcgaagtcgtccgcgcgggtctgtccaaggcggaacgaagtaagagcgaagggaag
Db	61	tccggcgaggtgtccgcgcgggtttgtcgaagccggcCAAGATAGAGCAGAGGGGAGG
OY	121	aggacgaagccggctggcggaggaacaggtacgcgcctcggcgcgctgtcaaggctgcgcg
Db	121	AGGACGAGCGGctggcgGAGGAGAGGTAAGCGCTCGGCGCGCGGTCAGAGGTGCTCGCGG
OY	181	gaatcccgcgccgcgtcggcgcccaacgcgcctcgcggcggggttcaattcgcgcgtctatt
Db	181	GAATCGCCCGCGCGctggcgGCGCACCGCCTCGCGCGGGGTCATTTCGCGCTATT
OY	241	ccggcggaagccacgcgcgcgcgtctgtcctcttcaacgcgaagaatcacaaggcggtgg
Db	241	CGGCGGAGACCAcCGCGCGGCGCTTGCCTTTCAcGCCAGAAATCTCAAGCGCGTGG
OY	301	ggctgcctccagtagaattcatcacgctcttgctgcgcgcgcgccttcttttggcc
Db	301	GGTGGCGTCCCGATGAGATTCTATCAAGTTTGGTGGCGCGCGCGCTTTTtttGGCGC
OY	361	tgcgaattcaagttctgtacttggggcaaatgtctgcagagataaggttgacgaagagttccc
Db	361	TGCATTATTAAgTTTGTACTTGGGCGCAAAAGCTGCAAGATAGGTGACCAAGAGGTTCC
OY	421	cttgcacccctgatatgaatcggaaccgggaacgtctgtgcacgtcttcaatcgaaagcgagctg
Db	421	CTTGACCCCTATGATGAATCGGAGCGGGAACGTGTGGCATGTCTTCTCAAGCGCAGCTG
OY	481	cacaaacatgcttcaagtgtaacaggttcgaacggcaaccttgcctccctaactcgcggaactac
Db	481	CACAACATGCTTTAGCGGTAACAAGTTCGAGCGGACCTTTGCTCTCATGCGGGGACCTAC
OY	541	cttgatgtcttccaagtctggtgtgtaaccttatgtcctaaggcaagtgtataagccgaaggagag
Db	541	CTTGATGTTTCCAATGCTCTGTGTGATCTTATTGCTAAAGCAGTGTAAAGCCGAGGGGAG
OY	601	tattgtgttccagcgcggttgtaacaatttcgaccccaatgatgctgaatcgaatccctctc
Db	601	TATGtGTTCCAGCGGCGGTGATCAATTCTGTGGCTTCAGATGGCTGACATGATCCCTCTT
OY	661	ccataacagacatttcaattgggaaggcgaactcaaccttaagatatcctaagaagacctg

Db	661	CCATATAGCACGTTTGATTGGAAAGCGACCTACTCTTAAGTATACCTCAAAAGACACTG	720
Qy	721	gtaataatagatgacattgacattgacgtggaattcagaaagcatgataccaagcatgtagaacat	780
Db	721	GTATATATAGATGACACTTGCGGTGATTCACGAAGCATGATTCAGCAAGTATGAACAT	780
Qy	781	ccgagatcattcatatggaagctgtgtcgaagcttgacatttgaagagcttgaggttaac	840
Db	781	CCGGGTACTTTCATTGGAGAGCTGTGTGGAAGCTTTCATTTTGAAAGACCTTGGAGTTAAT	840
Qy	841	tgatgtataataatgcccctcccaatgagttcaacagagctggagatactaaactctctccc	900
Db	841	TGATGTAAATTAATGCCCCCTCCCAATGATTCACAGACAGCTGAGTACTCAACTCTTCTTCC	900
Qy	901	aagaatgaactttgggagatacttacaacaactctcttccacaatgacagaatacacaa	960
Db	901	AAGATGAACCTTTGGGGATATTTCTACCATTAACCTCTTTTCACCAATACAAAGATACACA	960
Qy	961	tcaaggcggagataaaaaactgtgggcgtgatagccataaayagttcaaaaactttgtaga	1020
Db	961	TCAGGCGGGATAAAAAACTGTGGGCGGTGATGCCATTAATGATGATTAACAAACTTTGTAGA	1020
Qy	1021	gaagctcaaaaaaggggaattgagtgatccctggatgtgtcttcaaccatacagctggag	1080
Db	1021	GAGGCTCACAAACGGGGAAATTTGAGGTGATCCTGGAATGTGTCTTCAACCATACAGCTGAG	1080
Qy	1081	ggtataagaaatgttccaataatcataatcctaagggggctgcgaataactcaactacata	1140
Db	1081	GGTATGAGAAATGCTCCAAATATATATATATTAAGGGGCTGCATTAATACTATATATG	1140
Qy	1141	cttgacaccaagggagagatttaaaactctctgcgtgtgggaaatacctcaactgtaac	1200
Db	1141	CTTGACCCCAAGGAGAGATTTAATCTATTTGCGCTGGGAATACCTTCACACTGTAAT	1200
Qy	1201	catccctgtgtctgacaaattcatatgtagattgtttaagaatactgggtgaacggaaatgcat	1260
Db	1201	CATCCTGTGGTTCTCTCAATTCATTTGATGATTTGTAAGTAACTGGGTACACGAAATGCAT	1260
Qy	1261	gttataagtttctgttttgaatctctgatacacaagaacagaggtctcagactgtggagat	1320
Db	1261	GTTGATGGTTTTGCTTTTGATCTTTGCATTCATTAATGACCAAGAGTTCCAGTCTGTGGGAT	1320
Qy	1321	ccagatlaacgtgataagagatcccaataagaagtgacatgatacacaacagggacaactctc	1380
Db	1321	CCAGTAAACGTGATGAGAGCTCCAAATAGAAGTGAATGATTCACAACAGGGACACTCTT	1380
Qy	1381	gttaactccaaccaactatlgacatgatcaagcaatgacccaactctctggagcgctcaagctc	1440
Db	1381	GTTACTCCACCACTATTGTGACATGATCAGCAATGACCAACTCTTGGAGGCGCTCAACGCTC	1440
Qy	1441	attgcttgaagaatgaggtgtcgaagagccctcctaagaatgagccaattccctcaactggagat	1500
Db	1441	ATTGCTGAAGCATGGATGTCAGAGAGGCGCTCTTCAACAGTGAAGTCAATTCCTCTACGTGAT	1500
Qy	1501	gtttgctctgagtggaaatgggaagataccggagacattgtgcgtcaattcatataaagcaact	1560
Db	1501	GTTTGTCTGTAGTGGAAATGGGAAGTACCGGGACATTTGTGCGTCAATTCATTAAGGCACT	1560
Qy	1561	gatgagattgctgtgtgttttgcggaatgcttctgtgaaagctcacaacctaaccaaggtta	1620
Db	1561	GATGATTTGCTGTGTGTTTGGCGAATGCTTTGTGGAAATGCCACACTTATACAGGATA	1620
Qy	1621	agttgtggcaaatctgtttaaataagatcttgatggaatgaccccgagtttttataataacc	1680
Db	1621	AGTTGTGGCAATCTTGTAAATAGATTGATGATGATTCACCGGATTTTATATATATACC	1680
Qy	1681	acatgatatacacatacctaataataacaacatcatagtgatacgaatacttggtctaa	1740
Db	1681	ACATGATGATACACATCTTAATATATTAACAATATATGTTATGCAATATGCAATTTGGCTAA	1740
Qy	1741	gaagatcatgagtatacacaatagtgataataagttttaaacacccaacttgcgaatgaa	1800



481 cacaacatgcttaacggtacaggttcagagcaaccttgcctcactcgcgagactac 540  
|||||  
481 CACAACATGCTTACGGGATACAGGATTCAGAGGACCTTCTCCTCACTCGGGACACTAC 540  
541 cttgaagcttccaatgctcgtcgtcgaacccctacgtcaagcagtgaaacccagggag 600  
|||||  
541 CTTGAAGCTTCCAAATGCTGGTGGATCCCTTATGCTAAGGACAGATTAACCGAGGGAG 600  
601 tatggtctccagcgcgtggttaacaatctgctgcctcagaatgctgagcgtacccctc 660  
|||||  
601 TATGCTTCCAGCGCGCTGGTAAACAATTTGCTGGCTCAGATGCTGGCATGATCCCTCT 660  
661 ccaataagcagcttgatctgggaagcgagacctaccctctaagatacctccaagagacctg 720  
|||||  
661 CCTAATAGCAGCTTTGTTGGGAAGGCGACCTACCTTAAGATATCCCTCAAAAAGAGACTG 720  
721 gtaataatagagaatgcctgcgtggtgaatcaagaagcatgataccaagcaatgtgaacat 780  
|||||  
721 GTAAATATAGAGATGCCACTTGCCTGGATTCACGAAGCATGATTAACCAATGTGAACAT 780  
781 ccggtgacttcaatlgagcgtcgtcgaagccttgactatctgaagagcgttgagtaat 840  
|||||  
781 CCGGTACTTCAATTTGAGAGCTGTGTGACAGCTTACTATTTGAAGAGAGCTTGAGTTAAT 840  
841 tgcattgaaatlaatgcctcgcgtcgaatgatacgaagcgtgagatcaacctctctcc 900  
|||||  
841 TGTATTGAATTAATGCTTCCATGAGTTCAACAGCTGAGTACTCAACCTCTCTTCC 900  
901 aagaataaactttgggagatactcaacaataaactcttccaacaagaagaataaca 960  
|||||  
901 AAGATAACTTTTGGGATTTCTACCAATTAATCTTTTTCACCAATGACAAAGTATACCA 960  
961 tcaagcgaggaataaanaactggtgagcgatgagccaataagatcaaaactcttgtaaga 1020  
|||||  
961 TCAGGGGGGATTAANAACCTGAGGGCGATGCCATTAATGATGATCAAACTTTTGTAGA 1020  
1021 gaagctccacaacggggaatctgaagtgatccctgagatgttctccaacatacagctgag 1080  
|||||  
1021 GAGGCTCACAAACGGGGAAATGAGGTGATCCTGATGTGCTTCAACCAATACAGCTGAG 1080  
1081 ggttaataagagaatggtccaataatcaatcttaagggggtgcgaataactaactaata 1140  
|||||  
1081 GGTAAATAGAGATGCTCAATTAATTAATTAAGGGGGTGCATTAATTAATTAATTAATG 1140  
1141 ctgcaccccaaggagaggttctaactatctgcgtgaggaatlacctcaactgtaat 1200  
|||||  
1141 CTTCACCCCAAGGAGAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200  
1201 calccctggtctgcataatcaatctgtagatgtttaaagatacgtggtgagcgaaatgcat 1260  
|||||  
1201 CAACTGCTGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
1261 gttgagatggttctggttctggtctgcaataatgagagaggttccagctctgagat 1320  
|||||  
1261 GTTGAGATGTTTCTGTTTGAATCTTGCAATTAATTAATTAATTAATTAATTAATTAAT 1320  
1321 ccaagttaaagtgtaagagatccaatagaaggtgacatgatacacaacagagacacctt 1380  
|||||  
1321 CCAAGTTAAGCTGTATGAGAGCTCCAAATGAAGGTGACATGATCAACAAGGAGACACTT 1380  
1381 gttactccacacttaatgtaataatgaaacaaatctcttgagagcgtaacgtc 1440  
|||||  
1381 GTTACTCCACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440  
1441 attgcgtaaagcagtgagatgagagagcctctatacaagtaagtaagtaagtaagtaag 1500  
|||||  
1441 ATTGCTGAAGCATGAGATGAGAGGCTCTTATGAAGTGAATTAATTAATTAATTAATTAAT 1500  
1501 gttcgtcgtatgagaaatgagaaatgacgggagaaatctgctgcaatctcaataaagagc 1560  
|||||  
1501 GTTTCGTGATGATGAGATGAGAAATACCGGAGCAATTTGTGCTCAATTAATTAAGGCACT 1560

1561 gatgattctgctggtgttctgcccgaatgctcttgtagaagtcacacactaaccaggtga 1620  
|||||  
1561 GATGATTCTGCTGSGTGGTTTGGCCGAATGCTTTGTGGAAGTCCACACACTTACCAGGTA 1620  
1621 agtctgagcaactctgtaaaatgagctgagtgaaatgtaacccctgagatcttataatacc 1680  
|||||  
1621 AGTCTGAGCAACTCTGTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1680  
1681 acatgataatcaacatactaaataatacaacatacagtgatgataatgataatgataat 1740  
|||||  
1681 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1741 gaaatattagtgataacacagtgatgataatgaggtttaaaccacacactgagcaatgaa 1800  
|||||  
1741 GAAATATTAGTGTATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
1801 ggaacataagagctctcagatcatalcttaatctgctccggtgaaataacacactgaaat 1860  
|||||  
1801 GGAACATAGGCTTTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860  
1861 tcaagcaatgcatcttcttaagggggaagaacataatgataatgataatgataatgataat 1920  
|||||  
1861 TCAGACCATGTCATTTTAAAGGGGAGAGAAACATTAATTAATTAATTAATTAATTAATTAAT 1920  
1921 aagccatctcgaatctcaatgagtgatgctctctcgtgaaagaaagaaagactcat 1980  
|||||  
1921 AAGCCATCTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980  
1981 acctctacagtgatcaactcagatcagatctgataatctgaaagatgaaatgaaatgaa 2040  
|||||  
1981 ACTTCTATGCTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040  
2041 ttgtcgaataatgataatgataatgataatgataatgataatgataatgataatgataat 2100  
|||||  
2041 TTGTGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100  
2101 caagtaactctgattctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2160  
|||||  
2101 CAGTAACCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160  
2161 ctgtcgaagatcaactctgataatgataatgataatgataatgataatgataatgataat 2220  
|||||  
2161 CTGTGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
2221 cataataaagaatgataatgataatgataatgataatgataatgataatgataatgataat 2280  
|||||  
2221 CATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280  
2281 atcttaagctgaaatctggtgaggaagaagaatctcgaagatgtctgtcaaaagaatgaa 2340  
|||||  
2281 ATCTTAAGCTGAAATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2341 ggaagagcagatgagcagatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400  
|||||  
2341 GGAAGAGCAGATGCGCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2400  
2401 tctacatggtgataatgataatgataatgataatgataatgataatgataatgataatgata 2460  
|||||  
2401 TCTACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
2461 attcttaagctaaatctctgt 2520  
|||||  
2461 ATTCTTAAGTGAATTAATTTTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2521 tctgtcgtcctatgacacacactccgcaagaggtcgaaggtctgtggtcttgagagcttc 2580  
|||||  
2521 TCTGCTGCTCTATACCAATTTCCGCAAGAGAGTGCAGAGGCTTGTGCTTGAAGACTTTC 2580  
2581 caacggccaaacggctgcagatgagatgataatgataatgataatgataatgataatgataat 2640  
|||||  
2581 CAACGGCCAAACGGCTGCAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
2641 atagcgaatctgctgtccttccatgaaagatgaaagcagggagagatctatgtgctct 2700  
|||||



Db	2641	ATACCCGATTTCGTGCCCTTTTCCATGAANAAGTGAAGAAGCAGGGGAGATCATATGGCGCT	2700
Qy	2701	tcaacaccagcaccacttaacggccgcttggtgaatccccagagcgcgaggcgcggtgg	2760
Db	2701	TCAACACACGACGACATTACCGGGCCGTTGAGTCCAGAGCGGCAGGGCGCGTGCG	2760
Qy	2761	aaccggtggvtgaaacagagcaagcagacacatacgaatttcctaacgaagtactactg	2820
Db	2761	AACCGGTGGTGGACACAGGCAAGCCAGCACCATACGACTTCTCCACGACACTTACTCG	2820
Qy	2821	atgcgagctcaaccaataaccagttctgcgaatttctctactccaacttaaccacatgc	2880
Db	2821	ATCGGCGCTCAACCAATAACGAGTTTCGCATTTCTCTACTCCAACCTCAACCCCATGC	2880
Qy	2881	tcagctactcatcaggtcacccctagatctgcgcgccctgatglttgagaagacaatatata	2940
Db	2881	TCAGCTACTCATACGGGTCAATCCTAGATATGCGCCCTGAGTGTGAGAGACCAATATATACA	2940
Qy	2941	gttaaatatagtctcatatgctaataaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	2997
Db	2941	GTAATAATATAGTCTATATGTAAAAAIAAAAAAAAAAAAAAAAAAAAAAAAAAA	2997
RESULT	3		
LOCUS	TAE307689	2997 bp	mRNA linear PLN 17-DEC-2001
DEFINITION	Triticum aestivum partial precursor RNA for isoamylase (iso-1b gene).		
ACCESSION	AJ307689		
VERSION	AJ307689.1	GI:14331017	
KEYWORDS	iso-1b gene; Isoamylase.		
SOURCE	bread wheat.		
ORGANISM	Triticum aestivum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Poideae; Triticeae; Triticum.		
REFERENCE	1 (sites)		
AUTHORS	Genschel,U., Abel,G., Loerz,H. and Luetticke,S.		
TITLE	The sugary-type isoamylases in wheat		
JOURNAL	Unpublished		
REFERENCE	2 (sites)		
AUTHORS	Abel,G., Loerz,H. and Luetticke,S.		
JOURNAL	The Sugary-type Isoamylase In Wheat		
REFERENCE	3 (bases 1 to 2997)		
AUTHORS	Luetticke,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-JUN-2001) Luetticke S., Institut fuer Allgemeine Botanik, Universitaet Hamburg, Ohnhorststrasse 18, Hamburg, 20609, GERMANY		
COMMENT	Related entry AJ301647.		
FEATURES	Location/Qualifiers		
source	1..2997		
	/organism="Triticum aestivum"		
	/variety="Florida"		
	/db_xref="taxon:4565"		
	/tissue_type="developing kernel"		
	<1..2924		
	/gene="iso-1b"		
	<1..297		
	/gene="iso-1b"		
	/number=1		
	join(<1..297,398..1618,2146..2924)		
precursor_RNA	/gene="iso-1b"		
	1..2924		
exon	/gene="iso-1b"		
	join(<1..297,398..1618,2146..2924)		
mRNA	/gene="iso-1b"		
	join(<1..297,398..1618,2146..2924)		
gene	/codon_start=3		
	/product="isoamylase"		
	protein_id="CAC41016.2"		
CDS	/db_xref="GI:I7932913"		

[illegible]

Qy	601	taaggcttccagcgcggtgttaacaaatgtcctgcgtccagaatggcgtgacatgattccctctc	660
Db	601	tatggatgtttccacggcggtgtgtaaacatttctctggccttcagatggcctggacatgattccctt	660
Qy	661	ccatctagcaacgttttaatttgggaagcgacccctaccccttaagatatctctcaaaagacctg	720
Db	661	ccatattagcacgttttgatttgggaagcgacccctaccccttaagatatctctcaaaagacctg	720
Qy	721	gtcaataatagatgcacttctgttgaatttccacgaagcatgatctcaagcaatgtagaacat	780
Db	721	gtaataatattagatgacacttctgttgaatttccacgaagcatgatctcaagcaatgtagaacat	780
Qy	781	ccgggtacttcaatctggagctgtgttcgaagcttgactatttgaagagccttggagttaat	840
Db	781	ccgggtacttcttatttggagcctgtgttcgaagcttgactatttgaagagccttggagttaat	840
Qy	841	tgatatgaatttaagtcctccgtcccatgtatgttccaagaacctggagtaacttcaacctctctcc	900
Db	841	tgtatttgaaatttaagtcctccgtcccatgtatgttccaagaacctggagtaacttcttcttcc	900
Qy	901	aagaatgaacttttggggatcttccatctacaataaccttcttccacaaatgaagaatacaca	960
Db	901	aagatgaacttttggggatcttccatctacaataaccttcttccacaaatgaagaatacaca	960
Qy	961	lcaaggcgggataaaaaacactgtggcggtgaatgcataaataagtttcaaaaacttttgttaa	1020
Db	961	tcagcggcgataaaaaacactgtggcggtgaatgcataaataagtttcaaaaacttttgttaa	1020
Qy	1021	gaagctcacaaacgggggaattgaagtgtatccctgtgattgttcttccaacacatacagcttg	1080
Db	1021	gaagcgtcacaaaacgggggaattgaagtgtatccctgtgattgttcttccaacacatacagctg	1080
Qy	1081	ggtaatggaaatggttccaaattatatacttaaggggtgtcaataactacaatactaatg	1140
Db	1081	gctaattgaaatggttccaaattatatacttaaggggtgtcaataactacaatactaatg	1140
Qy	1141	cttgcacccaaaggagagtttctaactaacttccggctgtgtggaaatacctccaactgtaat	1200
Db	1141	cttgcacccaaaggagagtttctaactaacttccggctgtgtggaaatacctccaactgtaat	1200
Qy	1201	catcctgttggttcgttcaattcatctgtagaatgttctaagaatacttgggtgacggaaatgcat	1260
Db	1201	catcctgttggttcgttcaattcatctgtagaatgttctaagaatacttgggtgacggaaatgcat	1260
Qy	1261	gttgaatgttttcgttttgaactctgtacatccataatgaacgaaggttccagttcgttggat	1320
Db	1261	gttgaatgttttcgttttgaactctgtacatccataatgaacgaaggttccagttcgttggat	1320
Qy	1321	ccagttaaactgttatagagctccaatagaaaggtgacaatgaatcaacaaggagcaacctct	1380
Db	1321	ccaagttaaacgttatatgagacctccaatagaaaggtgacaatgaatgaacaggagcaacctctt	1380
Qy	1381	gttactccaacacttaattgacatgaatcaacaatgacacacaaatctctggaggcgtcaagctc	1440
Db	1381	gttactccaacacttaattgacatgaatcaacaatgacacacaaatctctggaggcgtcaagctc	1440
Qy	1441	attcgtcgaagcaatgggaatgcagaaggcctctatcaagtaagttcaatctccctcaactggaat	1500
Db	1441	atttcgtcgaagcaatgggaatgcagaaggcctctatcaagtaagttcaatctccctcaactggaat	1500
Qy	1501	gtttgtctcgaatgtgaatgggaatgcagaaggcctctatcaagtaagttcaatctcaataaggaact	1560
Db	1501	gtttgtctcgaatgtgaatgggaatgcagaaggcctctatcaagtaagttcaatctcaataaggaact	1560
Qy	1561	gatggaatttgcgtgtgttcttgcgaatgttcttctgtgaaagtcacacacataccaagta	1620
Db	1561	gatggaatttgcgtgtgttcttgcgaatgttcttctgtgaaagtcacacacataccaagta	1620
Qy	1621	agttctgtgcaataacttgttaatgagtttgaattgaatgttcaaccttcttataataaac	1680
Db	1621	agttctgtgcaataacttgttaatgagtttgaattgaatgttcaaccttcttataataataaac	1680

QY	1661	acatgatgatcaacatctcaaatatataaacaatctcaagtgatctgcatgacattggccaa	1740
Db	1661	ACATGATGATCAACATCTCAAAATATATATTAACAACATAGTGTATGCAATTTGGCTPA	1740
QY	1741	gaagatctagtgatacaactagtgctataatagaattctaaaccaactctgccaatga	1800
Db	1741	GAAGATATAGTGTAACACTAGTGTATATATAGTTTAAACCCCAACTTCCCAATGA	1800
QY	1801	ggaacataaggagcctctcagttatcttatcttatcttgccggtgaataatccactgaaaaa	1860
Db	1801	GGAACATATGGGGCTTCTAGTATCTTATTTATTTTCCGGTAAATATTCACATGAAAT	1860
QY	1861	tcacagccatgtcatctttcttaaggggggagaagaacataatgtatcttgcccccataag	1920
Db	1861	TCCACCCATGTCTATTTTATAGGGGGGGAAGAAACTATATGATTTGCCCCCTAAAG	1920
QY	1921	aagccatctcaagaatctcagtagtaagtgcccttctctgtaaaagaagaagacgactcat	1980
Db	1921	AAGCCATCTCAGAATATCTAGSTAAAGTGCTTTCTGTAAGAAAGAAACAGACTTCAT	1980
QY	1981	acttcctacgcgtgcctaacttagcctcgatgtgatalcttgtaagatgaatgccaacttaat	2040
Db	1981	ACTTCTATTCGGTGTCTAATCTAGCTCGATGTATATTGTAAGATGAATGCCAAATTAAT	2040
QY	2041	ttgtcggatbaatttgaatctgttatctacaanaatcttatcttggtctctcagaaatcaaac	2100
Db	2041	TTGTGGATTAATTTTATGATCTGTATATTCACAAATTTCTATTTGGTTCTGTAGAAATCAAC	2100
QY	2101	cagtaacttgatatttggaaccttgcaactctctatctgataatcaagccaggaagaagaac	2160
Db	2101	CAGTAACTTGATTTATTTGGCACTGCACACTCTTATTTATTAATCAAGCAGAGAGAAAGAAAC	2160
QY	2161	cttggaacagatatacaactcttgatgtgcacatgaatgaaatctcaactgctgaattcgata	2220
Db	2161	CTTGGCACAGTATCAACTTTGTATGTGCACATGATGAAATTAACACTGCTGATTTGGTA	2220
QY	2221	cataataaagaagatacaatcttacaanaatlgggagaacaacagagaatgagaanaatcaaa	2280
Db	2221	CATATTAATTAAGAAATGACATTTTACCAAATGGGGAAACAAACAGATGAGAAATTCACA	2280
QY	2281	atcttagcttggaattctgtyggggagaagaagaatctcgcaagatctgctgtcaaaagaatgga	2340
Db	2281	ATCTTAGCTGGAATTTGTGGGGAGGAAGAAATTTGGCAAGATTTGCTGTCAAAATATGA	2340
QY	2341	ggaagaaggcagatgctgcgaactctctctgtgtctcatabggtctcgaagaagtctcaatgt	2400
Db	2341	GGAAAGGCAGATGCGGCATTTCTTTGTTGTCCTATGTGTTCTTAAGAGATTTCCAAATG	2400
QY	2401	ctcacaatgggtgaatgaataatgcccacacaaagaagggaacacacacatacatatctgcata	2460
Db	2401	TCTACATGGGTGATGAATATATGTGCCACACAAAGGGGCCAACAAATCATATCTGCATNG	2460
QY	2461	attccttatgtcaattatcttctcgctvgggaataaaaaaagaacaabactcgaattgacgat	2520
Db	2461	ATTCTTATGTCAATTAATTTTCCGCTGGGATTAATAAAGAAACAATACACTGATGTGCACGAT	2520
QY	2521	tcctgtctgctcattgacccaatctccgcaaggagtgtgagaaggtcttgcccttgaggaacttc	2580
Db	2521	TCTGTGCTCTCATGACCAAAATTTCCCAAGAGATGCGAAGGCTTTGGCCCTGTAGGACTTTC	2580
QY	2581	caagcgccaagaatgtagtgagtgcaatgtcatcaacccctgggaagcctgatatgtctgaga	2640
Db	2581	CAAGCGCCAAGACGGGTGCGAGTGGATGTGATCAACCTGTGGAAAGCTGTATTGTGTAGA	2640
QY	2641	ataagccgatctgctgtgcctttctcatgaaagatgaaagaacagggcgagatatctatgtgcct	2700
Db	2641	ATAGCCGATCTGCTTCCCTTTCCATGAAGAAAGATGAAGACAGGCGAGATCTATGTGCGCT	2700
QY	2701	tcaaaacagacacacttaacggcgctgtgttgagatcccaagaacgagccaggaagcgccgttgg	2760
Db	2701	TCAAAACACGACACTTACCGGCGTGTGTGAGACTCCCAAGGCGCCAGAGGCGCCGCTGTGG	2760
QY	2761	aaccggtgtgtgacacaggaacgacagacacatacgaattccctcaacgagacttaactg	2820

```

|||||
Db 2761 AACCGGTGGAGACACAGGACACACACTTCCACACCACTTACCTG 2820
Qy 2821 atcgagctccacacacacacagcttcgcattccctactccaacctcaaccacgc 2880
Db 2821 ATGCGCTCTCACATACACAGAGTTCGCTTCTTACTCCAACTCTTACCCATGC 2880
Qy 2881 tcagctactacatcgatccatcgatctgcgcctgatcttgagagcaatataca 2940
Db 2881 TCAGCTACTCTCGCTCATCTCTATTTGCCCTGATGTTTGAGAGCAATATATACA 2940
Qy 2941 gtaataatactctatataataaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2997
Db 2941 GTAAATATATATGCTTATATGTAATAAAAAAAAAAAAAAAAAAAAAA 2997

RESULT 4
AF438328 2590 bp mRNA linear PLN 13-NOV-2001
LOCUS Triticum aestivum isoamylase precursor, mRNA, complete cds;
DEFINITION alternatively spliced.
ACCESSION AF438328
VERSION AF438328.1 GI:16905062
KEYWORDS bread wheat.
SOURCE Triticum aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 2590)
Neirphan,S., Khandelwal,R.L. and Chibbar,R.N.
Characterization of isoamylase transcripts produced during wheat
grain development and germination
Unpublished
2 (bases 1 to 2590)
Neirphan,S., Khandelwal,R.L. and Chibbar,R.N.
Direct Submission
Submitted (22-OCT-2001) Plant Biotechnology Institute, 110
Gymnasium Place, Saskatoon, SK S7N 0W9, Canada
Location/Qualifiers
1. 2590
/organism="Triticum aestivum"
/cultivar="Fielder"
/db_xref="taxon:4565"
/tissue_type="endosperm"
14. 2386
/EC_number="3.2.1.68"
/function="hydrolyzes alpha-1,6-glucosidic bonds of
starch, amylopectin, and glycogen"
/note="alternatively spliced; glycogen 6-glucohydrolase"
/codon_start=1
/product="isoamylase precursor"
/db_xref="GI:16905063"
/translacion="WTMAPEAPCLCARPSLAARRRSGSPARLRMRPNATAGK
GVEVCVAIVAAATKRVDEMEDEPAEDRALGACVILGMPPLGDTALAGVNF
AVYSGATTAALICFTPEPDLKADRTVEVPDLPLNRGNQHWVIEBELHMLGYR
FVGTGAPRCHGLDYSNVVDPAKAVISRGEYIPARKNNCPOMAGHPIPTSTED
WEGDLPRLYPQKDLVLYEMHLRGFTKHDSNVEHGTIGAVSKDLAKELGVNTEL
MPCHFNLELYSTSSKMFNGYSTINFPSPWRTYSGIKNGGDALNEFTFYREA
HKRGFEVLVDVFNHTAENNGPILSPRGVNTTYMLAPGEVYSGCNTFNCN
HPVVRQPIVDCLRWYTMREMHVDGFPREDLASTMTRGSLMDPVNVYGAPIEGMITEGT
PLVTPPLIDMISNDPILGCVKLIARMAPAGLYOGORPHRMVMSMNKKYDIYRQF
IKGTDFAGGAECLCGSPHYLOAGGRKPMHSINVCADGETLADLTYYKKYILPN
GENNNNNIYCHDSYVYFRMDKDOYSELHRCCLMTRFKRECEGLJEDPTAKRL
OMHGHPGKPMSENSRPAVAFSGMDROGEIYVARTSHLPAVVELPERAGRWBPVY
DMCKRPAVPDLTDLIDRALITHPSHFLYSMLYSSTIVLRPDV"
mat_peptide
/product="isoamylase"
2406..2411
polyA_signal 676 a 587 c 695 g 632 t
BASE COUNT
ORIGIN

```

```

Query Match. 56.2%; Score 1684.8; DB 8; Length 2590;
Best Local Similarity 78.8%; Pred. No. 4.3e-306;
Matches 2333; Conservative 0; Mismatches 2; Indels 627; Gaps 2;

1 ggtcgagccgagcgccgagcgcctcgagcaggttgcagccccaatgcagcgcgaggaaggg 60
90 ggtcgagccgagccgagcgcctcgagcaggttgcagccccaatgcagcgcgaggaaggg 149
61 tcgagagaggttgcagcgcgaggttgcagagcgaggaaggttagagagcgaggaaggg 120
130 TCAGCGAGGTGTCGCCCGCGGTTGTCAGAGCGCGCAGAGAGTAGAGAGAGGGGAGG 209
121 aggaagagcgaggttgcagaggaaggaaggaaggaaggaaggaaggaaggaaggaag 180
210 AGGAGAGCGCGGTGCGGAGGAGAGAGTACGCCCTCGCGGCGCGGTCAATTCGCCGTATTT 269
181 gaatgccgagcgagcgttgcagcagcagcagcagcagcagcagcagcagcagcagc 240
270 GAATGCCGCGCGCGGTGCGGAGGAGAGTACGCCCTCGCGGCGCGGTCAATTCGCCGTATTT 329
241 ccgagcgagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
330 CCGCGGAGCGAGCGCGCGCGCGCGCTCTGCTCTTACGCCAGAGATCTCAAGC----- 384
301 ggtcgctccgagtagaagttcaagccttcgtgcgcgcgcgcctttttggcc 360
385 ----- 384
361 tgcatttaagtttctacttgcaggaatgcaggaatgcaggaatgcaggaatgcaggaat 420
385 -----GGATAGGCTGACCGGAGGTTCC 409
421 ctgagcccgctgagtaagtcgagcagcagcagcagcagcagcagcagcagcagcagc 480
410 CTGAGCCCGCTGAGTATGAGGAGCGGAGAGGAGTGTGCAATGCTTCAATCGAGCGAGCTG 469
481 ccaaatatcttgcagcaggaaggttcgagcagcagcagcagcagcagcagcagcagcagc 540
470 CACAAATGCTTTAGCGGAGTACAGGTTGAGCGGAGCTTTGCTCTCACTCGGGCACTAC 529
541 ctgagtttccaaatgcgtggttgatcccttaagtaagcagtgtaagcaggaagggag 600
530 CTGAGTTTCCAAATGCTCGAGTGAATCTTATGTAAGCACTGATTAACCCAGCGGAG 589
601 tagtggtccagcgagtgtaagcagcagcagcagcagcagcagcagcagcagcagcagc 660
590 TAGTGTTTCCAAATGCTCGAGTGAATCTTATGTAAGCACTGATTAACCCAGCGGAG 649
661 ccaatagcagcgttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
650 CCATATAGCAGCTTTGATTTGGAGGAGCGAGCTTCACTTAAGATATCTCAAAAGAGCTG 709
721 gtaataatagatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780
710 GTAATATATGAGATGACCTTGCCTGATTCACAGAGCATGATTCAGCAATGTAGAACAT 769
781 ccgaggaacttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
770 CCGGGAATCTTCAATTTGAGAGCTGTGCAACCTTGACATATTTGAAGGAGCTTGAGTTAAT 829
841 tgcattgaatataagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
830 TGATTTGAATTAATGCCCTTCCATGATTAAGAGAGCTGAGATCTCAACTCTTCTTCC 889
901 aagatgaacttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
890 AAGATGAATCTTGGGATATTTACCATTAACCTTTTACCAATAGCAAGCATACACA 949
961 tcagcgaggaataaaaacgttgagcgtgtagtcataaataagagttcaaaactttgtaga 1020
950 TCAGCGCGGATTAATAAACTGTGGCGGTGATGCCATTAATAGATTCAAACTTTTGTAGA 1009

```

QY 1021 gaagctcaacaacggggaattgagtgatccctgagctgtcttcaaccatacagctgag 1080  
 |||||  
 Db 1010 GAGGCTCAACAACGGGGAATTGAGTGAATCGTGAATGTTGCTTCAACCAATACAGCTGAG 1069  
 QY 1081 ggttaataagataatggttcaataataataataataataataataataataataataata 1140  
 |||||  
 Db 1070 GGTAAATAGAAATGCTCAATATTTATTTAAGGGGCTGATTAATTAATTAATTAATTAAT 1129  
 QY 1141 ctgcacccaagagagagtttataactatctgcctgtggaatacctcaactgtaat 1200  
 |||||  
 Db 1130 CTTCGACCCCAAGGAGAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1169  
 QY 1201 catctgttggttcgtcaactcaatctgtgagatgttttaagatactggtgtgaagaaatgcat 1260  
 |||||  
 Db 1190 CATTCTGTGGTTCGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1249  
 QY 1261 gttgagtggttttggttttgatctgtgacccaataatgacaaaggttccagctcttgagat 1320  
 |||||  
 Db 1250 GTTGATGCTTTTCTTTTGTGATCTTGATCCATTAATTAATTAATTAATTAATTAATTAAT 1309  
 QY 1321 ccagctaacgttgatacggagctccaataagagtgacatgatacacaacagagacactctc 1380  
 |||||  
 Db 1310 CCAATTAACGCTGATGAGAGCTCCAAATAGAAAGTGAATGATCAACAGGAGACACTCTT 1369  
 QY 1381 gttactccacccaacttaattgacatacgaagaaatgaacccaattctgtgagagcgtaagctc 1440  
 |||||  
 Db 1370 GTTACTCCACCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1429  
 QY 1441 attgctgaagagatgagatgagagagagagagagagagagagagagagagagagagat 1500  
 |||||  
 Db 1430 ATTGCTTAAGATGAGATGAT 1489  
 QY 1501 gtttgcctgagtgagaaatgagaaatgacagagaaatgctgtcgtcaattcaataaagacat 1560  
 |||||  
 Db 1490 GTTTGGCTGAGTGAATGAT 1549  
 QY 1561 gatgagatctgctgagtggttttgcggaatgctctgtggaagtgccacacataaccagagta 1620  
 |||||  
 Db 1550 GATGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603  
 QY 1621 agtgtgtgcaatactctgttaaatgagtgagtgagtgacactgtgattttatataataacc 1680  
 |||||  
 Db 1604 ----- 1603  
 QY 1681 acatgatgatacacatctaataataataacatacagtagtatacagtaactgtgctaa 1740  
 |||||  
 Db 1604 ----- 1603  
 QY 1741 gaagtaattgatacacaatagtgctabalaataggtttaaacacccaacttgccaaatgaa 1800  
 |||||  
 Db 1604 ----- 1603  
 QY 1801 ggaacataagagctctctagtaatacttaatttgcgcggaataataccacactgtaaaat 1860  
 |||||  
 Db 1604 ----- 1603  
 QY 1861 tccagcaatgcatcttctagggggggaagaaactataatgtatttgcccccctaaag 1920  
 |||||  
 Db 1604 ----- 1603  
 QY 1921 aagccatctcagaattcagaatagtgatgtcttctctgtaagaagaaagaaacagactcat 1980  
 |||||  
 Db 1604 ----- 1603  
 QY 1981 acttctatcgtgtctaaacttagctcgaatgataattgttaagatgacaaatctaatt 2040  
 |||||  
 Db 1604 ----- 1603  
 QY 2041 ttgtcgaataattgatactgtatcacaataattctattgtttctctagaaatacaaac 2100  
 |||||  
 Db 1604 ----- 1603

QY 2101 cagtaactgttattgtgacacgcaactcttattgtatlaactcaaggcaggaagaaac 2160  
 |||||  
 Db 1604 ----- -CAGGACAGAGAGAGAGAAC 1622  
 QY 2161 ctgtgacagatcaactgtatgtgacataatgataatgataatgataatgataatgataatgata 2220  
 |||||  
 Db 1623 CTGGCAGAGATCAACTTTTATGTGACATGATGATGATGATGATGATGATGATGATGATGAT 1682  
 QY 2221 cataataaagaaatgacaaatttaacaaaatggggaagaaacagaatgagaaataca 2280  
 |||||  
 Db 1683 CATATAATAAGAAATACAAATTATACCAATGGGGAACAAACAGAGATGAGAAATACACA 1742  
 QY 2281 attctagctgaaatgtgtgggaagaaagaaatctgcagaatgtgtctcaaaagatgta 2340  
 |||||  
 Db 1743 ATCTTAAGCTGAATTTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1802  
 QY 2341 ggaagagcagatcgcaactctctgtgtctcactatgttctcagaagatctcagatg 2400  
 |||||  
 Db 1803 GGAAGAGCAGATGCGCAATTTCTTTGTTGTCATGTTTCTCAAGAGAGAGAGAGAGAT 1862  
 QY 2401 tctacatggttgatgataatgagcacaacaaagggggaacaaacatacactgtcagat 2460  
 |||||  
 Db 1863 TCTACATGAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1922  
 QY 2461 attctatgataatatttctgt 2520  
 |||||  
 Db 1923 ATCTTAATGATCAATTAATTTTGTGCTGGAGTAAATAAAGAAATTAATTAATTAATTAAT 1982  
 QY 2521 tctgtgtcctatgacaaatctccgaagagtgagagagtgctgtgactgtgagagcttc 2580  
 |||||  
 Db 1983 TCTGTGCTCATGATGACCAATTTCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2042  
 QY 2581 caacggccaacagcgtcagtgagtgatgatacagctgtggaagcctgagtgatgtgtgaga 2640  
 |||||  
 Db 2043 CAACGGCCAAAGGCTGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2102  
 QY 2641 atagcagatctgtgtgcttctccatgaagaatgaaagcagagagagagagagagagagag 2700  
 |||||  
 Db 2103 ATAGCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2162  
 QY 2701 tcaacacagcgaacttaacagcgt 2760  
 |||||  
 Db 2163 TCAACACAGCAGCAGCTTACCGGCGGTTGTGAGCTCCAGAGCGGCGGAGGCGGCGGCTGG 2222  
 QY 2761 aacgggtgtggaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820  
 |||||  
 Db 2223 AACGGGTGTGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2282  
 QY 2821 atgcgctctcaacatacagaatctgcgaattctctctctctctctctctctctctctctctct 2880  
 |||||  
 Db 2283 ATCGGCTCTCAACATPACACAGATGCTGCAATTTCTTACTCAACCTTACCCCATGC 2342  
 QY 2881 tcaagtaactcagctgatacctcagatgtgcgcctcagatgtgtgagagacaaatataca 2940  
 |||||  
 Db 2343 TCAGCTACTCATGCGTCAATCTCAGATTTGCCCTGATGTTTGAAGACCAATATATACA 2402  
 QY 2941 gtaataataatgtctataatgta 2962  
 |||||  
 Db 2403 GTAATAATATGCTATATATGTA 2424

RESULT 5  
 TAE301647  
 LOCUS TAE301647 2490 bp mRNA linear PLN 17-DEC-2001  
 DEFINITION Triticum aestivum mRNA for isomylase (1s01 gene).  
 ACCESSION AJ301647  
 VERSION AJ301647.1 GI:17932897  
 KEYWORDS 1s01 gene; isomylase.  
 SOURCE bread wheat.  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticaceae; Triticum.

REFERENCE	1 (bases 1 to 2490)
AUTHORS	Genschel U.
TITLE	Direct Submission
JOURNAL	Submitted (22-NOV-2000) Genschel U., Institut fuer Allgemeine Botanik, AMP II, Universitaet Hamburg, Ohnhorststrasse 18, Hamburg, 22609, Germany
REMARK	revised by submitter 22-MAR-2001
COMMENT	Related entry AJ307689
FEATURES	
source	Location/Qualifiers 1..2490
	/organism="Triticum aestivum"
	/cultivar="Florida"
	/db_xref="taxon:4565"
	/rname="developing kernel"
transit_peptide	6..214
	/gene="iso1"
CDS	6..2363
	/gene="iso1"
	/standard_name="sugary-1 isoamylase"
	/codon_start=1
	/evidence="experimental"
	/product="isoamylase"
	/protein_id="CAC82925.1"
	/db_xref="GI:17932898"
	/translation="MAKAPICAPRSLAARTRRPGPAPPLRRMRPNATAGKGVGEV CAVYEVATRADEGEDEDEPAEDRYALGACRYLACMPEDLQATALGAVNPVYSG GATAAALCLFPEDLKADRYTEVEPLDPLMKRTGNVHVFLEGLQMLGYRDTGTF APHGCHLDVSNVVDYPAKAVISRGEGVAPARNQNPOMAGIPLVSTFDEGDL PLRPORLDLVYEMHLRGFTKHDNSNEHPEPTIGAVSKLDYLLELVNCEILMPHE FNELEYSTSSKKNMGVSTINFSPTMYSGIKKCGRAINEFTFYEAHKGRI EVLIDVYENHTEKENGNEVSTLISFGVDNTYTYMALPAEGEFYNSGCTFNCHPVVR QFTDCLRWYTEMHVDGFRDLASIMTRGSLDVPVNVGAPLEGMTITGPTLMP PLINISNDYPLIGCVKLYAEKMDAGLYQNGQPFPHMVSFMNKKYDIYQPIKDD FGAGFAECLTGSFPHLTQAGRKPMHSINFCADHGFLLADLVYNNKYNLPGENNR DGEHNLSVYNGEGEGERARLSVLRKROMNPFVCLAVSGVMEFYMGDEYGHTKGSG NNNYTCHDSYVYFPMWCKEQYSDLHRCCLMTFRKCEGLDEPTAERLQMHGH QPGKFDMSNENRFVAFESMKIDROGEITYVAFNTSHLPAVELPERTGRMRBEVDTGKRP APVEDIDDLPERALTLTHQSFPLNSNLYPMLSYSVILVLRPY"
gene	6..2363
	/gene="iso1"
mat_peptide	215..2360
	/gene="iso1"
	/product="isoamylase"
BASE COUNT	619 a 573 c 683 g 615 t
ORIGIN	
Query Match	53.9%; Score 1614.8; DB 8; Length 2490;
Best Local Similarity	77.5%; Pred. No. 5.9e-293;
Matches 2296; Conservative	0; Mismatches 37; Indels 629; Gaps 3;
OY	1 ggtcggggcgccgcccgcgcctctcgacggttgcgaccacaaatcgacacggcggggaaagg 60
DB	67 GCGCGGGCGCGCGCGCGCGCTGCGACGCTGCGCGCCCAATCGACGCGCGGGAAGGGG 126
OY	61 tcggggagagtgccgcccggggtcttcgcaaggcggaagtaagagtgaggggggaag 120
DB	127 TCGGGGAGGCTCTGCCCGCGGCTTTCGAGAGTGTGACGACGAGGCCGAGGTGAGGGGAGG 186
OY	121 aaggacgagccggtcgccgagaggaacagtaacgagctcgccggcgccggttcgacgg 180
DB	187 AAGACGAGCGGGTTCGGCGAGAGGACAGGTACGCGCTCGCGCGCGGGGTCAATTCGCGCTACT 246
OY	181 gaatgcccgccgcgccttgaggcgcccaacgagctcgccggcggggttcaatttcgcgcctatt 240
DB	247 GAATGCCCGACGCGCGCTGCGGCGCACCGCGCGCTCGCGCGCGGGGTCAATTCGCGCTACT 306
OY	241 ccggagcagcaaccgcccggcgccctctgcctcttcaagcagaagaatcgaaggcggtg 300
DB	307 CCGGCGGACCCACAGCCCGCGCGCTTGCCTCTTCACGCCAGAAATCTCAAGC----- 361
OY	301 ggttcgctcccgagtagatltcaacgcttgcgtgcgcgcgcgccttttttgagcc 360

[illegible]

QY	1441	atctgcgtgaagcatalgagatgcgaaggagcctctactcaagtagtgcacattccctcacttgaa	1500
Db	1407	gtttcgtcaacacatgagatgcagagagccctctatcaaatgagtcattccctcacttgaa	1466
QY	1501	gttcgtctcgtagcgtgaatgagagatccggagacattgctgtcatttcaatttaaaagcact	1560
Db	1467	gttttgctctagcgtgaatgagagatccggagacattgctgtcatttcaatttaaaagcact	1526
QY	1561	gagagatctgcgtgtgttttttgcgcgaatgctctcttgtagaagtcacacacttaccaagta	1620
Db	1527	gattgattttctctgtgtgtttttgcgcgaatgctctcttgtagaagtcacacacttaccaagta	1580
QY	1621	agtcgtgcgaatactgtgaatgaatgagtggaatgctaccttgaaattttatataacc	1680
Db	1581	-----	1580
QY	1681	acatgtgatatacacatctcaaatatataaacatcaatagtgatgcatatgcatattggtctaa	1740
Db	1581	-----	1580
QY	1741	gaagtatttagtgaatacacagtgctataatagtttttaaacaccacacttgccaatgaa	1800
Db	1581	-----	1580
QY	1801	ggaacataagagccttctcgtatctatatttatttgcggtgaataacccactgaaaaat	1860
Db	1581	-----	1580
QY	1861	tccagcagatgcatttttttaggggggagagaacataatgatttgcctccctaaag	1920
Db	1581	-----	1580
QY	1921	aagcacatccaagaattcatatagaatgaaatgttccttcgtaaagaagaagaaacgacttcat	1980
Db	1581	-----	1580
QY	1981	acttctcatcgtgtcctaactagctcgaatgatatatttgaatgaatgaatgcacaaatttcat	2040
Db	1581	-----	1580
QY	2041	ctgctcgataaattgcatctgttatatcaacaattcttatttggcttctctagaaatcaaac	2100
Db	1581	-----	1580
QY	2101	cagtaactgttatactgacactgcaactctctattgtataatcgaacagaggaagaagaac	2160
Db	1581	-----	1580
QY	2161	cttgcaacagatcatcaactctgttatgtgcacatgatgatttcaactggtcgtatgtgttaa	2220
Db	1600	cttgcgcacagtatcAACCTTTGTATGTGTGCACACAGATGGATTTACTGGCTGATTTGGTAA	1659
QY	2221	catataaataagaatatacaatttccsaatgtyggagaaacaagagatgaganaaatcaaca	2280
Db	1660	catataaataagaatatacaatttccsaatgtyggagaaacaagagatgaganaaatcaaca	1719
QY	2281	atcttaactgtaaatgttgtyggagagagaagaatttgcgaagatgtgtctgtcaaaagatgaa	2340
Db	1720	atcttaactgtaaatgttgtyggagagagaagaatttgcgaagatgtgtctgtcaaaagatgaa	1779
QY	2341	ggaagagagcagatgacgaacatttctgttctgtctcaatgttcttcaagaagtcttcaatgt	2400
Db	1780	ggaagagagcagatgacgaacatttctgttctgtctcaatgttcttcaagaagtcttcaatgt	1839
QY	2401	tctacatggtgtgaatataatgctccacaacaagaaggggacaacaatacatatctgcaatg	2460
Db	1840	tctacatggtgtgaatataatgctccacaacaagaaggggacaacaatacatatctgcaatg	1839
QY	2461	attcttctgtcaaatatttctgcgttggaataaaagaacaatactcgtagtgttcttcccat	2520
Db	1900	attcttctgtcaaatatttctgcgttggaataaaagaacaatactcgtagtgttcttcccat	1939

OY	2521	tcctctgcctccatbaccacaatctccgaagaagatgcagaggtctctgacctbgaagacttc	2580
Db	1960	TCCTGTGCTCATATACCAAAATCCGCAAGAGTCGAGAGGCTTGCGCTTGAGCATTTTC	2019
OY	2561	caacgycaccaacgycctcgagatgacatgltcacgacctggaagccatgatctgtctgaga	2640
Db	2020	CAACGGCGCAACGGCTGCAGTGGCATGGTCTATCAGACCGCGGAACGCTATTTGGTCTGAGA	2079
OY	2641	atagccgattctgttgccctttccatbgaagtgtgaagaacagggggagatctcatitgscct	2700
Db	2080	ATAGCCGATTTGTTGGCTCTTTTCATGAAAGATGAAAGACAGGCGAGATCTATGTGGCT	2139
OY	2701	tcaaacaccagccacttlaacgycgctltgtltgacttcccaagcagcgcagagcgcgctgtgg	2760
Db	2140	TCAACACGAGCGACTTACCCGGCCGTTGTTGAGCTCCGGACGGCACAGGGCGCGGTGG	2199
OY	2761	aaccggtgtgtgacacaggaagccagcacatcagacttccctcaacgagacttactctg	2820
Db	2200	AACCGGTGTGTGGAACAACAAGCAAGCCAGCCATACGATTCCTCACTACACACTTACCTG	2259
OY	2821	atcgagctctcacatbaccacagttctcgcaatttcctcttacttccaactctcacccatgc	2880
Db	2260	ATCCGCGCTCTACCATACACAGTTCTCTCATTTTCTTAATCCAACTCTACCCCAATGC	2319
OY	2881	tcaagctactcatcgtgtcatctctaglatitgscgcccctgabtlttgagaagccaatatabataca	2940
Db	2320	TCAGCTACTCATTCGGTCACTCTAGATATGGCCGCTGATGTTTGAGAGGC - GGATATACA	2377
OY	2941	gtaaataatatgtctatataatga	2962
Db	2378	GTAATATATATGTATATATGTA	2399

RESULT	6
LOCUS	AX010486
DEFINITION	AX010486 2437 bp DNA linear PAT 06-SEP-2000
ACCESSION	Sequence 6 from Patent WO9598690.
VERSION	AX010486
KEYWORDS	AX010486.1 GI:997329.
SOURCE	bread wheat.
ORGANISM	Triticum aestivum
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticeae: Triticum.
AUTHORS	1 (bases 1 to 2437)
TITLE	Luetlicke,S., Abel,G., Loerz,H. and Genschel,U. Nucleic acid molecules which code for enzymes derived from wheat and which are involved in the synthesis of starch Patent: WO 9958690-A 6 18-NOV-1999:
JOURNAL	LUETTICKE STEPHANIE (DE); ABEL GERNOT (DE); LOERZ HORST (DE); GENSCHEL ULRICH (DE); HOECHST SCHERING AGREVO GMBH (DE)
FEATURES	location/Qualifiers
source	1..2437
	/organism="Triticum aestivum"
	/cultivar="Florida"
	/db_xref="taxon:4565"
	<16..2304
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC07515.1"
	/db_xref="GI:997330"
	/translation="PAPRLRRWRPNATAGRGVGEVCAAYEVATKADEGEDEEPVAE DRYLAGGACRYLAGMPPLGATLACGVNFAVYSGCATAAALCLTFPEDKADRYTEE VPLDPLNRRTGNWVHFIEGLDMLKSPFGDGFAPHCGHYLDVSNVVPYAKAVIY SRGEGYGPARGNQCMPOMAGMHPLEPYSTPEWEGDLPARYPKDKLYIEMHLRGTFRD SSNVEHPTGTIGAVSKUDYIKEIGVNCIEIAPCEHEVELEISTSSKMTFWGSIISF ESMPRTTSGGINKNGSDAINEKFTVEAKHRIEIVLDYVNHRAENGESITLSE GVNNTYTYMLADPGEFVYANGSCGFENECNHPVROPITVCLRWYEMHVDSEFPL ASIMTRSSLLMDPYNVYGADIEGDMITGTPLPPLIDMISNDPILGKVLVAEAD AGCLYXGCGFRHNMVSENNKKRYDIROPITKIGDFRAGGFAETICSPILYQAGCRK PMHSINIVYCAHDEGTADLVYTNKIVLPNGENNRDGEKNHLSNCGEEDSEFARLSV RLRRQRNRNFVCLMWSQGVPMFGDEYGTGKGNNTYCHDSYNYFPMWCKEDQYS







OY	238	atlcggygagacaccccgcgcgctctgcctcttcaegcagaagatctcaagcg	297
Db	413	ACTCGACGCGAGCACCGCGCGCCTCTGCCTCTTCACGACGAAAGATCTCAAGGC	470
OY	298	tggggtctgcctcccggtagagttcaatcagcttgcgtgcgcgcgccttttgg	357
Db	471	-----	470
OY	358	gccctgcaattgaattctgtactcgggcaaatgctcgcaagataaggtgacaggaaggt	417
Db	471	-----GGATAGGGGTGAGCGAGAGAGT	492
OY	418	ccccctgacccccctgtaatcgagccgggaacgctgtgcaagtctctcatcogaagcgag	477
Db	493	CCCCCTTGACCCCCCTGATGTAATCGAATCGGAGCGTGGCATGCTTCTCTCACTGGGGCAC	552
OY	478	cgcacacaaatctttaagggtacaaaggttcgacggaccccttgcctcaactcgsggac	537
Db	553	CTGCACGGCATGCTTATAGGTGACAGGTTGAGGGCACCTTCGCTCTCACTGGGGCAC	612
OY	538	tacccttgatcttccaaatgctcgctgtgtaaccttaatgctlaagacagatgaacgcgaag	597
Db	613	TACTTTGATGCTTTCGAATGTTGGTGGAGATCTTATGCTTAAGGACAGTAAAGCGGGAG	672
OY	588	gagtaatgtgtccacgcgctggtaacaatctgcctgcctcaagatgctgcaatgccct	657
Db	673	GAGTATGCTGTTCCGGCGCATGGTAACAATTGCTGGCCTCAGATGCTGCGATGATCCCT	732
OY	658	cttccaaatagaagctttgatctgggaaggagacctcccttaagatctctcaaaagac	717
Db	733	CTTCCATATAGCACGTTTGTGATTTGGGAAGGGGACCTTACTTAAATATCCCTCAAAAGAC	792
OY	718	ctgtgaataatagaatgcaactcgtcgtgtaatccagaaagcatgaatcaagaatgtga	777
Db	793	CTTGTATATATATAGATATGCACTTGCGGTGATTCACGAAGCATATTTCAAGCAATGTGAA	852
OY	778	catccgggtacttcatctgagactgtgtcgaaacttgactaattgaaaggacttgagct	837
Db	853	CATCCGGGACTTTCATTGGGGCGTGTGTCCAAAGCTTGACTTATTTGAAGGAGCCTTGAGT	912
OY	838	aatttgtatgaattaaatgcccctgcacatgattcaagaagcttgagatctcaactcttct	897
Db	913	AATTGATATGAATTAATGACCTTCACATGCAAGATTCACAGATGGAGTATGCAACCTCTCT	972
OY	898	tccaaagtgaactcttggggataatctacacataaactctcttcaaccaatgacaaatgac	957
Db	973	TCCAAATATAACTTTGGGGATTTCTACATATAACTCTTTTACCAATGACAGATAC	1032
OY	958	acatcagcgcyggaataaaaaactgtgggcgtgtgatgccaataagatctcaaaacttgt	1017
Db	1033	ACGTCAGGTGGGATTAATAAACTGTGGCGCGATGGGCAATAACGAGTTCAAAACCTTTGTA	1092
OY	1018	agaagaggtcccaaaacggggaaattgggggtgatactctgaaatgtgtcttcaacacataagct	1077
Db	1093	AAGAGATCTCACAAACGGGGAATTTGAGGTATCTCGATGTGTCTTCAACCATATACGT	1152
OY	1078	gaaggtgaatgagaatgttccaatattatcttaagggggtcgataataactataactat	1137
Db	1153	GAGGGTAATGAGAAATGCTCGATATATATCTTTAAGGGGCTGATATATACATACTACTAT	1212
OY	1138	atgcttgaccccaaggagagatttlaaactatctgctcgtctgpgaaataccttaactgt	1197
Db	1213	ATGCTTGACCCCAAGGAGAGTTTTATTAATATCTTGCGTGGGAAATACCTTCAACGT	1272
OY	1198	aattcaatcccggtgtctgtaactcatctgtatgttttlaagatactctgggtgacggaatg	1257
Db	1273	AATCATCCCTGTGGTGTCTCAATTTATTTGATTTGATTTAAGATATCTGAGATGGAATG	1332
OY	1258	catgttgatgggttctgcgttctgtactctgtgatacgaataagccagaagttccagctgtg	1317
Db	1333	CATATTTGATGGTTTGTGTTTGTATCTTTCGATTCCTAATATGACAGAGGTTTCAAGTCTGTGG	1392
OY	1318	gatccagttaaagctgtgataagagctccaatagaaggtgacatgtaacacaaagggacact	1377

Db	1333	GATCCAAATTAAAGGTATGGAAGCTCCAAATGAAGGTGACATGATCCAAACAGGACACCT	1452
Qy	1378	ctctgaactccaccactaattgaacatgaatcaagaaatgacccaattcttggaggcgtaag	1437
Db	1453	CTTGTTACTCCACACACTTATTGACATGATGACGAATGACCACAATTCTTGGAGGCTCAAG	1512
Qy	1438	ctcaatgctgaagcaatgggaltgcagaagggcctctcaagtaagtagtcaattccctcaatcg	1497
Db	1513	CTCATTCCTGAAAGCATGGATGGATGCGAGGCGCTCTATCAAGTATGATTCCTCCATCTGG	1572
Qy	1438	aatgtttggctctgtagtggaaatgggaagtaaccgggaatctgtgcataatcaataaaggc	1557
Db	1573	AATGTTGGCTGTGATGGAAATGGGAATACCGGACATTTGCTGCATCTTATTAAGGC	1632
Qy	1558	actgtagtaattctgtctgtgtttcttgcagaaatgctctcttggaaatccacaaccataacag	1617
Db	1633	ACTGATGGAATTTGGTGGTGGATTTTGGCGAATGTCCTTTGGGAAGTCCACAACTATAC---	1689
Qy	1618	glaagttgtggcaatctctglaaatgaagttagtgaaatgacacctgagatltttatataat	1677
Db	1690	-----	1689
Qy	1678	accacaatgatacatcacatcaataataacaacataatgatatgacataatgacattggc	1737
Db	1690	-----	1689
Qy	1738	taagaagatattgatactaacactagtgctataatagttttaacacccaacttgcacat	1797
Db	1690	-----	1688
Qy	1798	gaaggacaatagggctctctagtaatacttaattattgttcggtgaaataatccactgaa	1857
Db	1690	-----	1689
Qy	1858	aattccagcgaatgcatlttttllaggggggaagaagaactaatgatttgcctccctcaa	1917
Db	1690	-----	1689
Qy	1918	aagaagcatctcgaatctcaatagtaagtgtcttcttcgtaaagaagaacagact	1977
Db	1690	-----	1688
Qy	1978	catacttctatcgtctgaactagctgaatgataatttgaagaatgaatgcacaatt	2037
Db	1690	-----	1689
Qy	2038	aattgtcgaataattgatactgatacttcaataattcatttgttctctagaatca	2097
Db	1690	-----	1689
Qy	2098	aaccgaatactgttatattgtgcacatgcacaactctatgataatccagcaggagaagga	2157
Db	1690	-----CAAGCAGAGGAGAAAGA	1705
Qy	2158	aaccttgcacagatcaacttltgtaatgacatgaatgaattcaacactgcgtctgttgg	2217
Db	1706	AACCTTGGCACAGATCAATAATTTTGTATGTCACATGATGGATTATACATGCGCTGATTTGG	1765
Qy	2218	taacataataaagaagtacaattctaccaaaatggggagaacaacagagaatggagaatat	2277
Db	1766	TAACTATTAATAACCAAGATACAAATTTACCAATGGGAGGACATCAGAGATGCAAAATTC	1839
Qy	2278	acaactgaagtgtgaatttgggggaagaagaagaatctgcagaatgtgtctcaaaagaat	2337
Db	1836	ACAATCTTAGCTGAACCTGTGGGAGGAAGCAATTCGACAGATGCTAGTCAAAAGAT	1885
Qy	2338	tgaagaaagacagatgcgaattctcttctgtctcaatgttcttcaagagatgccaa	2397
Db	1886	TGAGGAAGAAGCAATGCGCAATTTCTTCTGCTCAATGTTTCTCAAGAGATTCCAA	1945
Qy	2398	tgcttctacatgggtgataatgataatggccaacaaaggggcaacaaataataatgccc	2457





QY 901 aagatgaactttgggagatattccataaactctttcccaatgacaagatacaca 960  
 Db 978 AAGATGAACCTTTGGGATATTCTACCATTAACCTTTTCACCAATGACAAGATACACA 1037  
 QY 961 tcagcgaggaataaaactctgagcgatgcataaataagatgacaaacttctgaaga 1020  
 Db 1038 TCAGCGGAGGATAAAACTGTGGCGGTGATGCCATAATGATGCTCAAAACTTTGTAGA 1097  
 QY 1021 gagctacaaacggggaattgagtgatccctgagtgctgtctcaacatacagctgag 1080  
 Db 1098 GAGCTCACAAACGGGGAATTGAGTGATCCTGGATGTTGCTTCAACACATACAGCTGAG 1157  
 QY 1081 ggtataagagaatggttccataattctcttaaggggtgcataatactactataatg 1140  
 Db 1158 GGTAAATGAGATGCTCAATATTATTCATTAGGGGGTCGATTAATCTACTACTATATG 1217  
 QY 1141 ctgcaccccaaggaggtttataactatctctgctggtggaatacctcaactgtaat 1200  
 Db 1218 CTGCAACCCAGGAGAGTTTATTAATCTCTGCTGTGGGATTAACCTTCACTGTAAT 1277  
 QY 1201 catctgtgtgtcgtcaatctatctatgtatgttttaagaactggtgagcgaatgcat 1260  
 Db 1278 CATCTGTGTGCTGCTCAATTCATGTAGATGTTAAGATCTGGTGACGGAATGCAT 1337  
 QY 1261 gttgagtgcttctgttctgtatctgcataatgacacaggttccagctctgtgagat 1320  
 Db 1338 GTTGATGCTTCTGCTTGTGATCTTCATCCATTAATGACAGAGGTTCAGTCTGTGGAT 1397  
 QY 1321 ccagtaacgtgtatgagctccataaga 1349  
 Db 1398 CCAATTACCTGTATGAGCTCCATGAGA 1426

RESULT 10  
 LOCUS AB015615 2288 bp mRNA linear PLN 06-JUL-2001  
 DEFINITION Oryza sativa mRNA for isoamylase, partial cds.  
 ACCESSION AB015615  
 VERSION AB015615.1 GI:3252793  
 KEYWORDS Isoamylase.  
 SOURCE Oryza sativa (cultivar: 'Fujihikari') early- to mid-milking stage endosperm tissue\_11b: endosperm mRNA, clone\_11b: lambda g11.  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Eriatoidae; Oryzaceae; Oryza.  
 1 (bases 1 to 2288)  
 Fujita, N., Kubo, A., Francisco, P. B. Jr., Nakakita, M., Harada, K., Minaka, N., and Nakamura, Y.  
 Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice  
 Planta 208 (2), 283-293 (1999)  
 9926133  
 2 (bases 1 to 2288)  
 Nakamura, Y.  
 Direct Submision  
 Submitted (17-JUN-1998) Yasunori Nakamura, National Institute of Agrobiological Resources, Department of Plant Physiology, 2-1-2 Kanonondai, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: nakaym@agr.affrc.go.jp, Tel: 81-298-38-8382, Fax: 81-298-38-8347)

FEATURES  
 source  
 1. .2288  
 Location/Qualifiers  
 /organism="Oryza sativa"  
 /cultivar="Fujihikari"  
 /db\_xref="taxon:4530"  
 /chromosome="8"  
 /map="determined"  
 /tissue\_type="endosperm"  
 /clone\_11b="lambda g11"  
 /dev\_stage="early- to mid-milking stage"  
 /tissue\_11b="endosperm"  
 <1. .2202

CDS

/codon\_start=1  
 /product="isoamylase"  
 /protein\_id="BAA29041.1"  
 /db\_xref="GI:3252794"  
 /translation="VEAVVMPERYALGACRVLAGMPAPLGAATLADGCVNFAVTSAGA  
 SAASLCFPPDLEADVEEYPLDPIFNRTGNVHWFIIEBELNMLYGVFDDMFAP  
 HGOYEDVSNVVDYPAKAVISGEVCPGCGDOMGMPIPLSPSTPMODLPL  
 RYPOKDLVTEMLRGFTKSSSVHEPRTYIGASIKSLDLKELGVNCEVLMPCHEFN  
 ELEYFSCSSNMNFWGYSYINFEFSPMIRYSSGGLINCGRDAINERTVREAHKQIEV  
 IMDVFNHTRAEKNGKPIISFRGIDNSTYMLAPKGFYNYSGGNTFNCNHPVRE  
 IVDCLRWYTEMHVDGFRFDLASIMTSGCSLMDPVNYGSPVEEDMTTGTPLATPL  
 IDMSINDPIGDVYKLAEMADAGLGYGVGEPHMKIWSNNGKXRDLYRQIKGTDF  
 AGCPAEKICGSPHLYOAGGRKPMHISINFCVAGCFEPLADLYTVYKXNNSGDNROG  
 ENHNSLNCEBEEGFACLSYRLRKRMKRNPFVSLATISQVPMYKQDEYHTNGGN  
 NTYCHDHYVYFEMDKKESSDLRFCSLMTKFKQCESLGLADFPQAQRLMHGHQP  
 GKPDSETSRFFVASTDETKELIYVAFNASHLPAVGLPERPGRMEPLVDTKAP  
 YDFLTDLDPRLVPSTSLISPSPISTPCSATPPSSINCSLMIERTYSIV"

BASE COUNT 587 a 508 c 593 g 600 t

ORIGIN

Query Match 36.7%; Score 1101.2; DB 8; Length 2288;  
 Best Local Similarity 68.1%; Pred. No. 1.5e-196;  
 Matches 1923; Conservative 0; Mismatches 273; Indels 628; Gaps 3;

QY 120 gagacgaagccggtgagcgaaggaacgaatgacgcctgagcgaggtgctgagc 179  
 Db 1 GTGAGGCGGCGTGATGCCGAGAGTACCGCGTGGCGGCTGCAGAGTCTCCGC 60  
 QY 180 ggaatgcccgcgcgtgagcgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 239  
 Db 61 GGAATGCCCGCGCGCTGGGGCCACCGCGCTCGACGGCGGGGTCAATTTCCGCTTAC 120  
 QY 240 tcagcgagcgaacgc 299  
 Db 121 TCCGCGCGGCGATCCGCGCGCTGCTGCTTACACCCCGAGAGATGTGAGGC---- 176  
 QY 300 gggctgcctccagtagatcagcattgctgctgcgcgcgcgcgcgcgcgcgcgcgc 359  
 Db 177 ----- 176  
 QY 360 ctgcacattgaattgttgaactgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 419  
 Db 177 -----GGATGAGGTGACTGAGAGCTTCC 200  
 QY 420 cctgacccctgataatgcgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 479  
 Db 201 GCTGATCCTCTGTTCAATCGGAGGGAATGTGCGACGCTTTCATCGAAGGAGACT 260  
 QY 480 gcaacaatgcttaagcgaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 539  
 Db 261 GCACAACATGCTGATACGGGTACAGGTTGATGATGTTCCGCCCTCACTCGCGCAGTA 320  
 QY 540 cctgagtgcttccaatgctgctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 599  
 Db 321 CTTCGATGCTTCATATGCTGCTGATGCTTATGCTTATGCTTATGCTTATGCTTATG 380  
 QY 600 gtatggttcccaagc 659  
 Db 381 GTATGCTTCCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440  
 QY 660 tccatataagcgttgaattggaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 719  
 Db 441 TCCGTACAGTACGTTGATGAGCAAGGAGCTACCTCTGAGATATCTCTCAAGAGACT 500  
 QY 720 ggtataatagatgacgc 779  
 Db 501 TGTATATCTATGATGCTTATGCTGCTTATGCTTATGCTTATGCTTATGCTTATGCT 560  
 QY 780 tcagtgacttccatgagctgtgcaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 839  
 Db 561 TCCAGGAGCTTACATTTGGGCTATATCAAACTTGACTATCTGAAGAGCTTGGAGTTAA 620

QY	840	ttgtatctgtaataaagccctgcacatgagttcaacgagcggagatctcaaccctcttc	899
Db	621	ctgtctaaagttgaaatgccctctgcacataaattcaatgacgtgagatcttcacgtctcttc	680
QY	900	caagaatgaacttttgggagatctctcaataaactctcttcaaccgaatgaatacac	959
Db	681	CAAGATGAACCTCTGTGGGATCTCCACGATAACTTTTTTTCACCATGTATGAATATTC	740
QY	960	atccagcgagataaanaaacctgtgggcgtgtagtcacataatgagttcaaaaactttgtaa	1019
Db	741	ATCAGSTGGGATAGAAAGAACTGTGGCGGTGATGCATTAATGAATGAAACTTTTGTTAG	800
QY	1020	agagagctcaacaaaggaggaattgaggtgatctctgagatgtgtcttcaaccatacagctga	1079
Db	801	AGAGGCTCTCAAAACGGGGAAATTGAGGTGATCATGAGATGTTGTCTCAATATCATAGCGCA	860
QY	1080	gggtatgaagaaatgtgcacataatacattaaagggatgcataataatacatacattat	1139
Db	861	GGGTATGAGAAAGAACCAATATTAATCATTTAGGGGATAGATATATGACATATCTATAT	920
QY	1140	gcttgcacacaaaggagagtttatataactcttcgctgtgggaaataccttcaactgttaa	1199
Db	921	GCTTGCCCCCTAAGGGAGATTTTACAATATATCTGGTGTGGGAATACCTTCAACGTGTA	980
QY	1200	tcaatccctgtgtctgcataatctcatctgttagatgtgttcaagaatactggtgacagaaatga	1259
Db	981	TCATCCCTGTGGTCCGTAATTTATTTAGATTTTAAAGTACTGGGTGACAGAAATGCA	1040
QY	1260	tgttgatggttttcgctttgagatcttcacataaagaacagaggttccagctcgtgtgga	1319
Db	1041	TTTTTATGATTTTCGTTTTGATTTTCATTCATTAATGACCAAGAGATGCAGTCTTTGGGA	1100
QY	1320	tccagttcaacgctgtatggagctccaaatagaaggctgacatgatacacaacagaggaacctct	1379
Db	1101	TCCAGTATATGTGTATAGGAAGTCCAGTAAAGAGTGACATACATACAGACAGGAACTCTCT	1160
QY	1380	tgttaatccacaactattatgacaatgatacagaatgacccaactctctggaggcgtcaagct	1439
Db	1161	TGCTACTCCACCACTTATTTAGCATGATACACATGATCCAAATCTTGGAGATCTCAAGCT	1220
QY	1440	catctctgaagcatgtagatgagagagagccctctcaagtgtgtcaatctccctcaatgtgaa	1499
Db	1221	CATTCCTGAGCATGGGATGGGGAGGCGCTCATCAAGTAGTCATAATTTCTCACTGGA	1280
QY	1500	tgttctgactgagatggaatgggagaatgacggagacatgtgtctcaattcatlaaaggcac	1559
Db	1281	AATTTTGCTAGAAATGCAATGGAAGAAGTACCGGACATTTGTCGTCAATTCAATTAAGGCAC	1340
QY	1560	tgaatgattctgctgtygtgttttgcggaatgtctcttgtgtgaagttcacacactataccaggt	1619
Db	1341	AGATGATTTTGGTGGTGGTTTCTCCTGAAATTTGTGTGGAAATCCACATCTATAC	1395
QY	1620	aagttgtggcaataactgttcaaatgagatgtgagtgtcaaccggaatlttttataataac	1679
Db	1396	-----	1395
QY	1680	caatgatgatcacacatctaaatatataacatcatagtgtagcatatgcatlttgcta	1739
Db	1396	-----	1395
QY	1740	agaagtattagtgatatcacatagtgcbatalatagaagttttaacacccaacttgccaatga	1799
Db	1396	-----	1395
QY	1800	aggaacataaggccttctcagttctctattattttgttcggtgtgaataatcaccatgaaaaa	1859
Db	1396	-----	1395
QY	1860	tccagccatgtcatctttttaagggggagaagaacataatgtattgtcccccctaaaa	1919
Db	1396	-----	1395
QY	1920	gaagccatctcagaatcatataggtlaagttgtcttctgtaaagaaagaacgacttca	1979

[illegible]

## RESULT 11

AX031277	LOCUS	AX031277	Sequence 16 from Patent WO9514314.	1706 bp	DNA	Linear	PAT 20-SEP-2000
AX031277	DEFINITION	AX031277					
AX031277	ACCESSION	AX031277					
AX031277.1	VERSION	AX031277.1	GI:10278607				
AX031277.1	KEYWORDS	AX031277.1	GI:10278607				
ORGANISM	SOURCE	ORGANISM	source				
Aegilops tauschii.	REFERENCE	Aegilops tauschii.					
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Aegilops.	AUTHORS	L.I.Z., Morell,M. and Rahman,S.					
1 (bases 1 to 1706)	TITLE	Regulation of gene expression in plants					
	JOURNAL	Patent: WO 9914314-A 16 25-Mar-1999;					
		GOODMAN FLEIDER LTD (AU) ; LI ZHONGYI (AU) ;					
		RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMM SCIENIT IND RES					
		ORG (AU) ; GROUPE LIMACRAIN PACIFIC PTY L (AU)					
FEATURES	FEATURES	FEATURES					
source	source	source					
		1..1706					
		/organism="Aegilops tauschii"					
		/db_xref="taxon:37682"					
		/tissue_type="ENDOSPERM"					
		<1..1599					
		/codon_start=1					
		/product="PARTIAL CDNA FOR HEXAPLOID WHEAT DBE"					
		/protein_id="CAC09652.1"					
		/db_xref="GI:10278608"					
		/translation="AVSKLDYLRKELGVNCTILMPCHEPNELEYSTSSKMNFGXSTII					
		NESFMRITRYSGGIRKNCGRDAINERKFTFRAHKRGLEVIIDVFNHTABGENGFPILL					
		DSFRVDNTTYTMLPKGEFNYSGCGNFENHPVYVQFIVDCIRYWMENHVDGFRFF					
		LFIATIRGSSILMDPENVYVYGAPEIGDMITGCTPLVPLTIDMISNDILGQVKLIAEA					
		MDAGILYOVQGFPHMNVSWENNGKRYRDIYRQFIRKGTGDFAGCELCGSHLQACGGA					
		RKPMHSINFCVADHGFTLGLDLYTNKKYNLDPNGNNRDNENHNSMNCGEGERARLS					
		VKRLARKRMRRCLFPEFCVMSQCVPMFPMGDEYGHKKGKNNNTYCHDSVYVYRPMKKKEQ					
		YSDLRHRELTAKTRKECEBEGLEDFPRAEPLQNHGHPGCPDMSSENRFYAESKDEQ					
		ROGEIYVAENFNTSHLPVAVLEPERAGRRPVPDGRKAPYDFELTDPLDRALTIHOPS					
		HELNSNLVPMHSYSSVILVLRPDY"					
BASE COUNT	467 a	355 c	415 g	469 t			
ORIGIN							
Query Match	35.8%;	Score 1074.4;	DB 6;	Length 1706;			
Best Local Similarity	74.9%;	Pred. No. 1.5e-191;					
Matches 1621;	Conservative	0;	Mismatches 16;	Indels 527;	Gaps 1;		
QY	799	gctgtgtgtgaagcttgcattcttgaagaagcttgaagttaattcatgtgaattaatgcc	858				
DB	1	gctgtgtgtgaagcttgcattcttgaagaagcttgaagttaattcatgtgaattaatgcc	60				
QY	859	tgcacatgaagcttgcattcttgaagaagcttgcattcttgcagaatgaactttgggga	918				
DB	61	tgcacatgaagcttgcattcttgaagaagcttgcattcttgcagaatgaactttgggga	120				
QY	919	tattctacataaactcttcttcacacaatgacagaatgacacatacacaatgcagcgcgga	978				
DB	121	tattctacataaactcttcttcacacaatgacagaatgacacatacacaatgcagcgcgga	180				
QY	979	tgtgtggcgtgtatgacataaagtgtcaaaacttttgttaagaagaagcttcaaaacgggga	1038				
DB	181	tgtgtggcgtgtatgacataaagtgtcaaaacttttgttaagaagaagcttcaaaacgggga	240				
QY	1039	attcgagtgatccctgtgatttgttcttcaacccaataagcctggaaggttaataagaatgtcca	1098				
DB	241	attcgagtgatccctgtgatttgttcttcaacccaataagcctggaaggttaataagaatgtcca	300				
QY	1099	ataatcatcttaagaagggtgcataaatacatatactatattgtctgcacccaaggagagag	1158				
DB	301	ataatcatcttaagaagggtgcataaatacatatactatattgtctgcacccaaggagagag	360				
QY	1159	ttttataactattctgtgctgtgtgggaataccttcaactgtaatcatccctgtgtgtgtcca	1218				
DB	361	ttttataactattctgtgctgtgtgggaataccttcaactgtaatcatccctgtgtgtgtcca	420				

OY	1219	ttctcttgatgattggtttaagaatctctggcgagcggaaatgcagtgtgattcggtttc	1278
Db	421	TTTCATTGTAGATTGTTTAAATACTGAGTGATGGAAATCCATGTGTGATGGTTTCGTTT	480
OY	1279	gactctgcattccaaatgacacgaaggctccagctctgagtcctgaatccagtcgatga	1338
Db	481	GATCTTGCAATCCAAATAGACACAGAGTTCCAGTCTGTGGGATCCAGTTAACTGTATGGA	540
OY	1339	gtcccaatagaaggctgacatgatcaacaacggagacccctctgttaactccaccattc	1398
Db	541	GCTCCAAATAGAAAGGTGACATGATCACACAGGGACACCTCTGTACTCCACACTTATTT	600
OY	1399	gaactgtcttcagcaatgtgcccgaattctcttgaggggtcacaagctcaattgtctgaagcaggat	1458
Db	601	GACATGATACGAATGTGACCCAAATCTTGGAGGGGTCAACCTCATTTGCTGAACATGGGAT	660
OY	1459	gcagagagccctctcaacagtaggccaattccctccactgagaaatggtctgctgagtgagat	1518
Db	661	GCAGGAGGCCCTCATCAATGATAGGTCAATTCCCTCACTGGAAATGTTTGGTCTGATGGAAAT	720
OY	1519	ggggaagtaccgggagacattgtgcgtcaattcaattcaataaagccatgataattgtctgtgt	1578
Db	721	GGGAAGTACCGGGACATTGTGGCCCAATTCATTTAAAGCATGATGAGATTTTCGTGTGT	780
OY	1579	tttcgcgaatgtctctgttggaagtcacacactaacacagtgtaagtgtgtgcacattcgt	1638
Db	781	TTTGGCAATGTCTTTTGTGGAACTCCACACTATAC-----	816
OY	1639	aaatgagltgagtgaaatgtccacctggaatttttatatatacacatgatgatcacactc	1698
Db	817	-----	816
OY	1699	aaatatataacaatcatagtgatgatcatatgcaattgtgcgaagaagtattgtatataca	1758
Db	817	-----	816
OY	1759	ctagtgctataataaggttttaacaccccaacttgcacaatgaaagacaatagggcttcta	1818
Db	817	-----	816
OY	1819	gttatctattattcttctccgggtgaataatccactgaaaaattccagccaatgtaatttc	1878
Db	817	-----	816
OY	1879	taggggggagagaacatactatgtattgtcccccctaagaagcactcagaatcca	1938
Db	817	-----	816
OY	1939	taggtgaagtgtcttctctgtaaagaaagaaaacgactcctaacttctactcgtgtctaa	1998
Db	817	-----	816
OY	1999	cttagctctcagatataattgttgaaatgaaagcacaattcaattgtgcgataattgtac	2058
Db	817	-----	816
OY	2059	tgatatacaaaattcttaattgtttctctagaatacaaacocgtaacttgttatgtgc	2118
Db	817	-----	816
OY	2119	actgcgaactctctatgatcaatcaagcgagggagaaagaaacctgtgcacagtatcaact	2178
Db	817	-----CAGCAAGAGAAAGAAACCTTGGCACTATACACT	853
OY	2179	ttgtatgtgcacatgatgagattcaaatgtgctgactctgtttaaataataaataagagataca	2238
Db	854	TTGTATGTGCACATGATGAGATTTTACACTGGGTGATTTTGTCATATATATAACAAGTACA	913
OY	2239	atttaaccaaatgtgggagacaacagagatgagagaanaatcaacaactcttagctgaaatgtgt	2298
Db	914	ATTTACCAAAATGGGAGAGACAAATAGAAATGGAAGAAATACAAATCTTACTGTGAATTTGTG	973



OY	2299	gggaaaggaaggaaatctgcgaagaatgtctgtcaaaaagtcttgaggaagggcagaatgacga	2358
Dd	974	GGGAGGAAGCAAAATTCGCAGAATGTCTGTCAAAGATTGGAGAGGGCAGATGCCCA	1033
OY	2359	attcttctgtttgttcatacgtgttctcaaggagttccaatgtctcacatcggtgtgat	2418
Dd	1034	ATTCTTCTTTGTTGTTCTCATNGTGTTTCTCAAGSAGTTCCAAATGTTTTACATGGCGATGAA	1093
OY	2419	atgaccacaacaaagggygaccaaacabaactatgcataatcttaatgtcaattt	2478
Dd	1094	ATGGCACACAAGAGGGGGCAACAACAAATACACTAGTCGATGATTTCTTAATGCAATTTAT	1153
OY	2479	tccgcttggtgataaaaaagaacaaatactctgagtctgacacgaatcttcgtgcctcatagacca	2538
Dd	1154	TTCGGTGGGATPAAAAAAGAACAAATACCTCTGACTTCCACAGATTCTGCTGCCATGACCA	1213
OY	2539	aattccgaagaagatgacgaaggctctgacctgtagagacttccaaacggccaacggctgc	2598
Dd	1214	AATTTCCGACAGAGATGBCAGAGGCTCTTGCGCTTGAGSAGCTTCCACAGGCCGCAAGCGCTGC	1273
OY	2599	agttgcatalgtcatcagcccttgysaagccatgatlgtlctgagaaatagccgaatcgttgtcc	2658
Dd	1274	ACTGGCANTGTCATCAGCCTGGGAAGCCTGATTTGGTCGAGAAATAGCCGATTCGTTGCCT	1333
OY	2659	ttccatatgaagaatgaagaacagggcagagatctatgtgtgacctcaacacagcaactaac	2718
Dd	1334	TTTTCATGAAAATGAAACAGACAGGCGAGANTCTAATGTGGCCTTCAACACACACCCTTAC	1393
OY	2719	cggcgctgtgttagctcccagaagcgcgacagcgccggtgggaacccggtgtgtgagacag	2778
Dd	1394	CGGCGCTGTTCAGCTCCACAGACGGCGACGGCGCGGTGGGAACCGGTGGTGGACACAG	1453
OY	2779	gcaagccagacacatacgaactctcccaaccgaagacttaactatgcgtccctcacacatac	2838
Dd	1454	GCAAGGCACACACATATGACTTCTCAGACGAGACTTACTGTGCGGCTCCACACATAC	1513
OY	2839	accagttctcgacattctcctctactccaaccttaecccaatgctacgactactaactgtgtca	2898
Dd	1514	ACCAGTTCCTCATTTCCCACTCACTCAACCTTACCCCCTGCTCAGCTACTCATGCGTCA	1573
OY	2899	tccatgatatggccccctgtgtgttgagagacaatatagagtaaataatgtctata	2958
Dd	1574	TCCGTGATATGGCCCCCTGATGTTGAGAGACAAATATATACAGTAAATATATGTTCTATA	1633
OY	2959	tgtca 2962 	
Dd	1634	TGTA 1637	
RESULT 12			
AR072143			
LOCUS	AR072143	2712 bp	DNA linear PAT 18-FEB-2000
DEFINITION	Sequence 1 from patent US 5912413.		
ACCESSION	AR072143		
VERSION	AR072143.1	GI:7223031	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2712)		
TITLE	Myers,A.M. and James,M.Graham.		
JOURNAL	Isolation of SUL1, a starch debranching enzyme, the product of the		
FEATURES	maize gene sugary1		
SOURCE	Patent: US 5912413-A 1 15-JUN-1999;		
location/Qualifiers			
1..2712			
BASE COUNT	636 a 646 c 745 g 685 t		
ORIGIN	/organism="unknown"		
Query Match	33.4%; Score 1000; DB 6; Length 2712;		
Best Local Similarity	65.7%; Pred.No. 1,5e-177;		

[illegible]

QY	1156	gaattatataactatctcgcgtctggaaatcactccaaactgaatcaatccctcgtgctcgt	1215
Db	1213	gagttttatattatattctggttggaatattccttaattgaatcaatccctgtaatccgt	1272
QY	1216	caattcactctgaattctttaaagatcctggctgacaggaatgcattgtatgcttccgt	1275
Db	1273	gaattttatagtgattgcttgagactctgagatctggaatcgaattgattgattgattgctg	1332
QY	1336	ggaagcccaatagaagctgacatgcatacaacagagacacccctcgttacccecaact	1395
Db	1393	cgaaagttccattgacagctgacatgatttaccagacagacaccccttctgcccacactt	1452
QY	1396	attgacatgacagcaattcctgtaggagcgtcaagctccattgctgaagacatg	1455
Db	1453	attgacatgattgacatgacatgacatgacatgacatgacatgacatgacatgacatg	1512
QY	1456	gagtcagagagccctcctacagatgacatgacatgacatgacatgacatgacatgacatg	1515
Db	1513	gattcagagagagctcttctatcagaagctgacatgacatgacatgacatgacatgacatg	1572
QY	1516	aattgagagacagcagacatctgctgacatgacatgacatgacatgacatgacatgacatg	1575
Db	1573	aattgagagagagctcttctatcagaagctgacatgacatgacatgacatgacatgacatg	1632
QY	1576	gatttgcagagagctcttctgtagaagctgacatgacatgacatgacatgacatgacatg	1635
Db	1633	gatttgcagagagctcttctgtagaagctgacatgacatgacatgacatgacatgacatg	1671
QY	1636	tgtaacagagctgtagaagctgacatgacatgacatgacatgacatgacatgacatgacatg	1695
Db	1672	tgtaacagagctgtagaagctgacatgacatgacatgacatgacatgacatgacatgacatg	1671
QY	1696	tcataataatacaacatcagatgctgacatgacatgacatgacatgacatgacatgacatg	1755
Db	1672	tcataataatacaacatcagatgctgacatgacatgacatgacatgacatgacatgacatg	1671
QY	1756	aacttgcagatgacatgacatgacatgacatgacatgacatgacatgacatgacatgacatg	1815
Db	1672	aacttgcagatgacatgacatgacatgacatgacatgacatgacatgacatgacatgacatg	1671
QY	1816	ctagttactatcttacttctcgcgtgacatgacatgacatgacatgacatgacatgacatg	1875
Db	1672	ctagttactatcttacttctcgcgtgacatgacatgacatgacatgacatgacatgacatg	1671
QY	1876	tttaagaggggagagaaactatctgatttgcctcctaaagaagccatctcagaat	1935
Db	1672	tttaagaggggagagaaactatctgatttgcctcctaaagaagccatctcagaat	1671
QY	1936	tcaatagtaagcttcttctgtaagaagaagaaacgcttcaacttctatctcgtgc	1985
Db	1672	tcaatagtaagcttcttctgtaagaagaagaaacgcttcaacttctatctcgtgc	1671
QY	1996	taacttagctcgatgatacttctgtaagaatgaaatgacaaattaaattgctcgataaattg	2055
Db	1672	taacttagctcgatgatacttctgtaagaatgaaatgacaaattaaattgctcgataaattg	1671
QY	2056	atcgtctatctcaaaatttcttatttctcctctagaatacacaacagtaactgttatt	2115
Db	1672	atcgtctatctcaaaatttcttatttctcctctagaatacacaacagtaactgttatt	1671
QY	2116	ggcagcgcacacttcttactgataatcagcagagagaaagaaaccttgcacagatca	2175
Db	1672	ggcagcgcacacttcttactgataatcagcagagagaaagaaaccttgcacagatca	1705
QY	2176	aacttgcagatgacatgacatgacatgacatgacatgacatgacatgacatgacatgacatg	2235
Db	1706	gctttgattgacacacagatgatttaccatgcttatttggcacaatcattacattacattacatt	1765

QY	2236	acaaattaccaaatgagggaacacacagagttgagaaataacaaattctgtcgtgaatt	2235
Db	1766	ACAACCTGTCAAAATGCTGAGGAGACTTCAGAGATGGGGAAAATCATATCTTACGTGGCAATT	1825
QY	2296	gtgggaggaagaagaaattctcgcaaatgtctgtcacaagaattgagaaagaagagatgc	2355
Db	1826	GTGGGAGGAAGAGAAATTTTCAAGTGTGTGACGCCGAAGATTAAAGGAAGGCAAAATGC	1885
QY	2356	gcaattccttgtttgtctcctcctcgaagtctccaaagtctcgaattctctacatggtgtg	2415
Db	1886	GCAATTTCTTGTGTGTGCTTATGTTTCTCAGGAGATGCCAATGTCTACATGGGCGCATG	1945
QY	2416	aatttgcacacaaagaaggaggagacaaacaaataactctgcacgatctctatccaatt	2475
Db	1946	AATATGGTCACACAAGGAGGGAACAACAATACGTACTGTCATGACCATTAATGTCAATT	2005
QY	2476	attctgcctgtgataaaaa--aagacacatactctgattgcacgaattctgcgtccctca	2532
Db	2006	ATTTCCTGTTGGGATTAAGAGAGAAACAACTCTGTGATTTTACAGATTTCGCCGTCA	2065
QY	2533	tgaaccaattctcgcaagaagtgcgaaggctcttgaccttgaaacttccaaaggccaatt	2592
Db	2066	TGACCGAATTCGGCAAGAAGTAATGTGAATCTTGGCCCTGAGAGACTCCCGACTTCAGAAC	2125
QY	2593	ggctgcagttgacatggttcatacagctctggaaagctgtatgtgtcgaagaatgccgattcg	2652
Db	2126	GGTTGAATGTGCAGGCTATACGCCGGGAAGCCTGTGATGTGTCAAGAGCAAGCCGATTGCG	2185
QY	2653	tgtgcctttcgaatgaagaatgaagaacagaggcagatctatgtgtgaccttcaaacacagcc	2712
Db	2186	TTGGCTTACCATGAAGAGACAAACCAAGCCGAGATTTACGTGGCTTTCACACCAATGC	2245
QY	2713	acttaacggcggctgtgttgagctcccaagacgacgagggcgcgggtgaggaacgggtgtg	2772
Db	2246	ACCTTCCGGTGGTGTGTGGCTTCCAGAGCGCTGTGGTTCCGATGGGAGCGGGTGGTG	2305
QY	2773	acacaggaagcagacacatacgaactctctcaaccgacgaactacatgacgcgtctca	2832
Db	2306	ACACCGGCAAGAGAGACCAATATGACTTTCACCGAGTGGCCTGCCAATATGCTGCTGCA	2365
QY	2833	ccataacacaaattctcgatattctctctctacacttcaaaccttaccatgtctagctactcat	2892
Db	2366	CCGCTTACCAATTTCTCTATTTCCTCAACTCCCAATCTCTATCTATGTCTAGCTACTCTCT	2425
QY	2893	cgatcatctagtattgtcgccctgatatgtttgagaaagacaaatatata	2938
Db	2426	CGATATCTTGTATGTGCGCCCTGTGTCTGAAGAAGACAGATAC	2471

RESULT	13			
LOCUS	ZMU18908	2712 bp	mRNA	linear
DEFINITION	Zea mays Sulp (Sugar)1		mRNA, partial cds.	PLN 06-MAR-1996
ACCESSION	U18908			
VERSION	U18908.1		GI:758725	
KEYWORDS	.			
SOURCE	maize.			
ORGANISM	Zea mays			
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 2712)			
TITLE	James, M.G., Robertson, D.S. and Myers, A.M. Characterization of the maize gene sugary1, a determinant of starch composition in kernels			
JOURNAL	Plant Cell 7 (4), 417-429 (1995)			
MEDLINE	9529095			
REFERENCE	2 (bases 1 to 2712)			
AUTHORS	Myers, A.M.			
TITLE	Direct Submision			
JOURNAL	Submitted (19-DEC-1994)			
	Alan M. Myers, Biochemistry and Biophysics, Iowa State University, 2110 Molecular Biology Building, Ames, IA 50011, USA			



Oy	1576	ggttttcggcaatgctcttttgtagaattccacacactatcacagtaagtcttgcgaataact	1635
Db	1633	gcttttgctgaatgccccttatcttggaagctccacactttatgc-----	1671
Oy	1636	tgtaaatgcttgcagagaaatgtcacctcgtgattctttatataccaatgatgatcaca	1695
Db	1672	-----	1671
Oy	1696	tcataaaatatacaaatcatagtgatgacatgacatgtgcatttgcctaagaagatcatagtgat	1755
Db	1672	-----	1671
Oy	1756	acacactagctctatatataagttttaacaccccacacttgcacatgaagaacataggcgctt	1815
Db	1672	-----	1671
Oy	1816	ctagtatctcttatcttatttgcgcgtgaaatcatccactgaanaaatccacagcatgtcat	1875
Db	1672	-----	1671
Oy	1876	ttcttaggggggagaaagaaactaatatgtatttgcgcccttaaaagaagacatctcaagat	1935
Db	1672	-----	1671
Oy	1936	tcaatgataagtcttcttctctgttaaaagaagaagaagacacttcatcttctctacgtgc	1995
Db	1672	-----	1671
Oy	1996	taactagctcgatgtatatttgtaaagatgacgtccaaatttaatttctgcgataatttgc	2055
Db	1672	-----	1671
Oy	2056	atctgtatcatcaacaattctatatttgcctctagaatatcaaacacgttaactgttatt	2115
Db	1672	-----	1671
Oy	2116	ggcaactgcacaactctcttaatgaattaacacagcagagagaagaagaaaccttgcacagatca	2175
Db	1672	-----CAGGCAAGGGGAGGAAGCCCTTGCCACAGTATCG	1705
Oy	2176	accttgtaigtgcacatgaatgaatttacaactgcgtgaatttgcgtacaatatagaagaat	2235
Db	1706	gcttttgctatgtccacacagatgattttacactgcctgattttggctcacatpacatattgcacagt	1755
Oy	2236	acaattcacacaatctggggagaaacaacaacagagatgtgagagaataatcaacacttaactgcgaaat	2295
Db	1766	ACAACtGTGCAAAATGCTGAAGGACTTCGAGAAATGGGGAAATCAATAATCTTAAGCTGGAATT	1825
Oy	2296	gtcggagaggaagagaatctgcagaagtgtctgtcaaaagaatvtgaagaagagcagatgc	2355
Db	1826	GTGGGAGAGAGAGAAATTTGCAAGCTGTGCAGTCCGAAAGATTAAAGGAAGAGCGCAAAATGCG	1885
Oy	2356	gcaattctcttgccttctcctcagaatgttctcgaagaagtccaatgttctcactgsgtgatg	2415
Db	1886	GCAAAATTCTTTGTTCTCTTATGTTTCTAGGGAGTTCCAAATGTTTACATGTTGGCGCATG	1945
Oy	2416	aaatatgcacacacaaaaggggcacaacacatatcatcttgcactgaattcttatatccaat	2475
Db	1946	AATATGSGTCAACAACAAAGGGAGGGAACAACAATAGTACTGCCATGACACATTATGTCAATT	2005
Oy	2476	attctgcctggaataaaa---aagaacaatactctgaagtgtcacacgaattctgtcgtccca	2522
Db	2006	ATTTCCGTTGGGATPAGCAAGGAAGAACAAATCCCTGTATTTGTACAGATTCGTGCCGTCTCA	2065
Oy	2533	tgaaccaaatctcgcagaagagtgcgaggggtcttgcgccttgaggaactttccaaagggccaaac	2592
Db	2066	TGACCAAAATTCGCCAAGAAGATGTGAATCTTTGGCCTTGAGGACTTCCCGCACTTCAGAAC	2125
Oy	2593	ggctcagatgtgcacatgtcatcaacgcttgcgaagcctgtgattgtctgaagaatagccgaattcg	2652
Db	2126	GGTGTAAATATGGCAACGGTCAATCAACCCGGGAAGCGTCACTGCTGTGAAGGAAGCAAGCGGATTCG	2185
Oy	2653	ctgccttctccatgaaagatgaaagacagaggcgagatctatgtgccttccaaacacagcc	2712

	Db	2186	TTCGCTTACCATGTAGAGACGAAACCAAGCGAGATGTACGTGGCCTTCAACACAGTC	2245
	Qy	2713	actiaacgagccgttcttgtagctcccagaacgcgcgaaggcgctggaaaccggttggtg	2772
	Db	2246	ACCTTCGCGTGTTGTTGGCTCCAGAGCGCTCTGGGTTCCGATGGAGCCGGTGGYG	2305
	Qy	2773	acacaagcaaacccaacatacgaactcttcacacgaacgaattaccgatgcgtctca	2832
	Db	2306	ACACGCCAAGAGAGCACCAATATGACTTCCTCACCGAAGGCGCTGCACATATGTCGTGA	2365
	Qy	2833	ccatcacaccagtcttcgcgattctccctacacccacacctaccatgctacgcatcat	2892
	Db	2366	CCGTTACCAGTCTCTCTATTCTCTCAACTCCAACTCTTAATCTCATGCTCAGCTACTCT	2425
	Qy	2893	cggtaactccatgatatgcgccctgatgtttgagaccaatatala	2938
	Db	2426	CCATCATCTCTGTATGTCGCCCTGATGTCTAAGAAGACAGATATCA	2471
	RESULT	14		
	LOCUS	AF142590	983 bp	mRNA linear
	DEFINITION	Triticum aestivum isoamylase 1 mRNA, partial cds.		PLN 12-AUG-1999
	ACCESSION	AF142590		
	VERSION	AF142590.1		GI:4929176
	KEYWORDS			
	SOURCE			
	ORGANISM			
		bread wheat. Triticum aestivum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum. 1 (bases 1 to 983) Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C. Analyses of isoamylase gene activity in wild-type barley indicate its involvement in starch synthesis Plant Mol. Biol. 40 (3), 431-443 (1999)		
	JOURNAL	MEDLINE	99364538	
	REFERENCE	PUBMED	10437827	
	AUTHORS		2 (bases 1 to 983) Sun,C., Sathish,P., Ahlandsberg,S. and Jansson,C. Direct Submission Submitted (13-APR-1999) Plant Biology, Genetic Center, The Swedish University of Agricultural Sciences, P.O. Box 7080, Uppsala S-75007, Sweden	
	JOURNAL			
	FEATURES			
	source		location/Qualifiers	
		1..983	/organism="Triticum aestivum"	
			/db_xref="taxon:4565"	
	CDS		<1..>983	
			/codon_start=3	
			/product="isoamylase 1"	
			/protein_id="AA033890.1"	
			/db_xref="GI:4929177"	
			/translation="DLPRTYRQKDLVYEHLNGFTKHDSNWEPHTFGIVASKLDY LKELEIEMPCHEPNELEYSTSSKNMFWGSTINFSPMTRYTSIGIKKGGRDA INEKTFVRREAHKRKGIIVILDVFNTHAEENENGPILSPGVNTTYMLAPGEFVN YSGGGFENHNHPVROFIYDCLEVVTYENHNVGFRFLASIMTRGSLLMDPVNYGA PIBEDMTTGTPRLPLIDMSNDPILGGVKLYLAEBMADGGGLYOGGPPHMVMWSM NGRKRDITRGPIKTDFPAGFPACELCSPHLYOVAGSRKRMHSINTVCAND"	
	BASE COUNT		271 a 188 c 241 g 283 t	
	ORIGIN			
	Query Match		31.1%; Score 932.8; DB 8; Length 983;	
	Best Local Similarity		98.7%; Pred. No. 5e-165;	
	Matches 940; Conservative		0; Mismatches 12; Indels 0; Gaps 0;	
	Qy	686	ggagactaccctctaagatatccccaagaagccctgtataataatagagatgcaacttgcgtg	745
	Db	1	GCGACCTACCTCTAAGATATCCTCAAAGGACGCTGGTAAATATAGATGATGACATTCGGCTG	60
	Qy	746	gattcaagaagcatgattcaagaagcaatgagaacatccgggtacttcaattcgtagcgttgt	805











QY 1381 gtatctcaccacttatatgacatgatacgaatgacccaattcttggagcgctcaagctc 1440  
 Db 1381 gtatctcaccacttatatgacatgatacgaatgacccaattcttggagcgctcaagctc 1440  
 QY 1441 attctgaagacatggatcgagcggaagcctctcaatgaagtcgaattccctcaactggat 1500  
 Db 1441 attctgaagacatggatcgagcggaagcctctcaatgaagtcgaattccctcaactggat 1500  
 QY 1501 gttctgctcgatgtagaactggaagtaaccggaacatgtgctgcaattcaattaaagcact 1560  
 Db 1501 gttctgctcgatgtagaactggaagtaaccggaacatgtgctgcaattcaattaaagcact 1560  
 QY 1561 gatvgatttgcctgctgcttcccgcaatgctcttggagagtcacacctaaccagta 1620  
 Db 1561 gatvgatttgcctgctgcttcccgcaatgctcttggagagtcacacctaaccagta 1620  
 QY 1621 agtctgagcaatctatgtaaaatgagttgagtgagtcacacccggattttatatatacc 1680  
 Db 1621 agtctgagcaatctatgtaaaatgagttgagtgagtcacacccggattttatatatacc 1680  
 QY 1681 aacatgatatcacacatccaataataaacaatcatagtgatgcatatgcatatgctcaa 1740  
 Db 1681 aacatgatatcacacatccaataataaacaatcatagtgatgcatatgcatatgctcaa 1740  
 QY 1741 gaagtaattagtgatatacactgctgcatatataagttttaacacccaacttgcgaatgaa 1800  
 Db 1741 gaagtaattagtgatatacactgctgcatatataagttttaacacccaacttgcgaatgaa 1800  
 QY 1801 ggaacacatagggctctctgatactcttatatttgcgcgtagaataaccactggaanaat 1860  
 Db 1801 ggaacacatagggctctctgatactcttatatttgcgcgtagaataaccactggaanaat 1860  
 QY 1861 tccagccatgcatctttcttaaggaggagggaagaacataatgatttgcctccccaagaag 1920  
 Db 1861 tccagccatgcatctttcttaaggaggagggaagaacataatgatttgcctccccaagaag 1920  
 QY 1921 aagccatctcagaattcatalaagtaagtctcttctctgaagaagaagaacgacttcat 1980  
 Db 1921 aagccatctcagaattcatalaagtaagtctcttctctgaagaagaagaacgacttcat 1980  
 QY 1981 acttctcatcggtgctaaactgaactgcgtatataatttgaagatgaagccaatttaac 2040  
 Db 1981 acttctcatcggtgctaaactgaactgcgtatataatttgaagatgaagccaatttaac 2040  
 QY 2041 ttgtcggataattgcatctgtatctcaacaatttctatctgttctctcagaacaac 2100  
 Db 2041 ttgtcggataattgcatctgtatctcaacaatttctatctgttctctcagaacaac 2100  
 QY 2101 cagtaactgttcatctgacatgcaactctctatgtattcaalcaggacaggaagaagaac 2160  
 Db 2101 cagtaactgttcatctgacatgcaactctctatgtattcaalcaggacaggaagaagaac 2160  
 QY 2161 ctgtgacacagatatacaactttgtatgtcgcacatgcatggaatttcaactgtgctgtaa 2220  
 Db 2161 ctgtgacacagatatacaactttgtatgtcgcacatggaatttcaactgtgctgtaa 2220  
 QY 2221 catataaagaagaatgatacaatttaacaatgaggagaacaacagagatggaagaatacaca 2280  
 Db 2221 catataaagaagaatgatacaatttaacaatgaggagaacaacagagatggaagaatacaca 2280  
 QY 2281 atcttaactggaatctgtgaggaggaagaagaatctgcgaagatgtctgtcaaaatgta 2340  
 Db 2281 atcttaactggaatctgtgaggaggaagaagaatctgcgaagatgtctgtcaaaatgta 2340  
 QY 2341 ggaagagcgagatcgcaacttcttctgtctcaatggttctcgaagagttccaatgt 2400  
 Db 2341 ggaagagcgagatcgcaacttcttctgtctcaatggttctcgaagagttccaatgt 2400  
 QY 2401 tctacatggtgtagaataatgacacacaaaggagggaacaacatactatctccatg 2460  
 Db 2401 tctacatggtgtagaataatgacacacaaaggagggaacaacatactatctccatg 2460

QY 2461 attctatgtcaattatcttctgctgctggaataaaagaagaacatactatgattgcaccgat 2520  
 Db 2461 attctatgtcaattatcttctgctgctggaataaaagaagaacatactatgattgcaccgat 2520  
 QY 2521 tctgtctgctcactgacacaaatctccgaaggaatgtgaggtcttgccttgaagacttc 2580  
 Db 2521 tctgtctgctcactgacacaaatctccgaaggaatgtgaggtcttgccttgaagacttc 2580  
 QY 2581 caacggccaacacgctgacatgtagcatatgcatcaagccttggagaagcctgattgtctaga 2640  
 Db 2581 caacggccaacacgctgacatgtagcatatgcatcaagccttggagaagcctgattgtctaga 2640  
 QY 2641 atagccgattctgtctcttccatggaagaatgaaagacagggcgagatctatgtgcct 2700  
 Db 2641 atagccgattctgtctcttccatggaagaatgaaagacagggcgagatctatgtgcct 2700  
 QY 2701 tcaacacacacacacttaccgagcggtgtgtgagctcccaagacgacgagggcgccgtggg 2760  
 Db 2701 tcaacacacacacacttaccgagcggtgtgtgagctcccaagacgacgagggcgccgtggg 2760  
 QY 2761 aacgggtgtggaacacaggaagcagacacatacgaacttctcaccgacgaacttaccctg 2820  
 Db 2761 aacgggtgtggaacacaggaagcagacacatacgaacttctcaccgacgaacttaccctg 2820  
 QY 2821 atcgcgctctcaccatatacagcagttctcgcatttctctactccaactctccaccatgc 2880  
 Db 2821 atcgcgctctcaccatatacagcagttctcgcatttctctactccaactctccaccatgc 2880  
 QY 2881 tcaagctacatcgctcatctccagatgtgcgcctgagtttggaggaacataataaca 2940  
 Db 2881 tcaagctacatcgctcatctccagatgtgcgcctgagtttggaggaacataataaca 2940  
 QY 2941 gtaataatattgtcatatgttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 2997  
 Db 2941 gtaataatattgtcatatgttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 2997

## RESULT 2

AAX34654  
 ID AAX34654 standard; cDNA; 1706 BP.

AAX34654;

05-JUL-1999 (first entry)

Partial cDNA for hexaploid wheat DBE.

Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBS;

starch branching enzyme; starch soluble synthase; debbranching enzyme;

endosperm; wheat; barley; granule-bound synthase; glutenin; starch;

grain softness protein I; bacterial isoamylase; glycogen synthase;

WSBE I-D4 gene; ss.

Triticum tauschii.

Key Location/Qualifiers  
 CDS 1..1704  
 FT /\*tag= a  
 FT /note= "partial DBE; the attr and stop codons are  
 not indicated"

W0914314-A1.

25-MAR-1999.

11-SEP-1998; 98WO-AU00743.

20-MAR-1998; 98AU-0002509.

12-SEP-1997; 97AU-0009108.

(GSTR ) COMMONWEALTH SCI & IND RES ORG.

(GOOD-) GOODMAN FIELDER LTD.

(LIMA-) GRP LIMRAIN PACIFIC PTY LTD.

PA (AUSU ) UNTV AUSTRALIAN NAT.  
 XX  
 PI Li Z, Morell M, Rahman S;  
 XX  
 DR WPI: 1999-229525/19.  
 XX P-PSDB: AAY06918.  
 PT New isolated cereal plant enzyme genes used for, e.g. expression of  
 XX antisense sequences of granule bound synthase  
 PS  
 XX Claim 47: Page 100-102; 171pp; English.

CC The invention relates to a novel enzyme of starch biosynthetic pathway  
 CC in a cereal plant, where the enzyme is selected from starch branching  
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching  
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
 CC SBE I of rice or maize. The methods and products can be used for  
 CC targeting expression specifically to the endosperm of the seeds of cereal  
 CC plants such as wheat or barley. They can be used for the expression of  
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,  
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
 CC can be used for modifying the characteristics of starch produced by a  
 CC plant. The present sequence represents the wheat DBE partial cDNA  
 CC sequence.  
 XX  
 XX Sequence 1706 BP; 467 A; 355 C; 415 G; 469 T; 0 other;

Query Match 35.8%; Score 1074.4; DB 20; Length 1706;

Best Local Similarity 74.9%; Pred. No. 4e-222; Indels 527; Gaps 1;

Matches 1621; Conservative 0; Mismatches 16;

QY 799 gctggttcgaagcttgaacttgaagagcttgaagtaattgatatgaattatgcc 858  
 DB 1 gctggttcgaagcttgaacttgaagagcttgaagtaattgatatgaattatgcc 60  
 QY 859 tgcgaagatcaacagagctgagctactcaactctctcccaagatgaaatttgggga 918  
 DB 61 tgcgaagatcaacagagctgagctactcaactctctcccaagatgaaatttgggga 120  
 QY 919 tatttaccataaactcttccacatagaagaatgacatacagcgagataaaac 978  
 DB 121 tatttaccataaactcttccacatagaagaatgacatacagcgagataaaac 180  
 QY 979 tctgggcgtatgcataaattgaattcaaaactttgttaagagagctccaaacgggga 1038  
 DB 181 tctgggcgtatgcataaattgaattcaaaactttgttaagagagctccaaacgggga 240  
 QY 1039 attgaggtgactcctgagatgttcttccaacatacagctgaaggtaattgaaatgtgcca 1098  
 DB 241 attgaggtgactcctgagatgttcttccaacatacagctgaaggtaattgaaatgtgcca 300  
 QY 1099 atttcttcatttaaggagctgagataactataactatgcttgcagccaaaggagag 1158  
 DB 301 atttcttcatttaaggagctgagataactataactatgcttgcagccaaaggagag 360  
 QY 1159 ttcttaaacacttcctgctctgggaataccttcaactgtaatacctctgtgttcgcca 1218  
 DB 361 ttcttaaacacttcctgctctgggaataccttcaactgtaatacctctgtgttcgcca 420  
 QY 1219 tctaatgttagatcttgaataactctgggtgaacgaatgcatgttgaatttccgtttc 1278  
 DB 421 tctaatgttagatcttgaataactctgggtgaacgaatgcatgttgaatttccgtttc 480  
 QY 1279 gatctgcacacataagacagaggttccagctctgaggatccagtaacgctgtatgga 1338  
 DB 481 gatctgcacacataagacagaggttccagctctgaggatccagtaacgctgtatgga 540  
 QY 1339 gctccaaatagaaggtgacatgatacaacagggacaccttctgttactccacacctatt 1398  
 DB 541 gctccaaatagaaggtgacatgatacaacagggacaccttctgttactccacacctatt 600

QY 1399 gacatgatacgaatgacacacacattcttggagggctcaagctcattgctgaagcattggat 1458  
 DB 601 gacatgatacgaatgacacacacattcttggagggctcaagctcattgctgaagcattggat 660  
 QY 1459 gcaagagctctatcaatgaatgaattcaattccctcactggaattgttgcgtgaatggaat 1518  
 DB 661 gcaagagctctatcaatgaatgaattcaattccctcactggaattgttgcgtgaatggaat 720  
 QY 1519 gggaaagacgggagacattgtgcgtcaattcaatgaagacactgtatgtctgtgtgt 1578  
 DB 721 gggaaagacgggagacattgtgcgtcaattcaatgaagacactgtatgtctgtgtgt 780  
 QY 1579 ttggcgaagctcttcttgggaagctcaccacattatgacaggaatgttctgcaaacattgt 1638  
 DB 781 ttggcgaagctcttcttgggaagctcaccacattatgacaggaatgttctgcaaacattgt 816  
 QY 1639 aaatgaattgagtgatgatacctggaatttttataatacacaatgatacacaact 1698  
 DB 817 ----- 816  
 QY 1699 aaatataaacaacataagtgatgatacattggttaagaagatagtgataca 1758  
 DB 817 ----- 816  
 QY 1759 ctagtgtatataatagttttaacacacacattgccaatgaaggacaataggtcttcta 1818  
 DB 817 ----- 816  
 QY 1819 gtatctattatttctgcgtgataatcaccatgaaataatccacgcatgatttt 1878  
 DB 817 ----- 816  
 QY 1879 taaggggggaagaagaacataatgatttgcctcccaataaagaagcattccgaatca 1938  
 DB 817 ----- 816  
 QY 1939 taagtaagtgcttcttctgtaagaagaagaagacttcaattctctacgtgtctaa 1998  
 DB 817 ----- 816  
 QY 1999 cttagctgataatatttgaatgataatgccaatttaatttgcgataattgtctc 2058  
 DB 817 ----- 816  
 QY 2059 tgtattacaacatttcaatttgcctctagaatacaaacagtaactgttattgctc 2118  
 DB 817 ----- 816  
 QY 2119 actgcaactcttattatcaatcagcgcaagaggaagaagaaacttgcgacagatcaact 2178  
 DB 817 ----- 816  
 QY 2179 ttgtatgtgcatgataatgatttaccactgctgttctgttaacataataagaagataca 2238  
 DB 854 ttgtatgtgcatgataatgatttaccactgctgttctgttaacataataagaagataca 913  
 QY 2239 atttacaacatggggagaacacaacagagatgggagaatacacaactcttagctggaattgtg 2298  
 DB 914 atttacaacatggggagaacacaacagagatgggagaatacacaactcttagctggaattgtg 973  
 QY 2299 gggaggaaggaatattcgcaagattgtctgcacaaagattgaaggaagaggaattgcga 2358  
 DB 974 gggaggaaggaatattcgcaagattgtctgcacaaagattgaaggaagaggaattgcga 1033  
 QY 2359 attccttcttctgctcattgcttctcaagagattcccaatgcttcaatggttgaat 2418  
 DB 1034 attccttcttctgctcattgcttctcaagagattcccaatgcttcaatggttgaat 1093  
 QY 2419 atggcacaacaaaaggggacaacaaatacactgcccattgcttcttattt 2478  
 DB 1094 atggcacaacaaaaggggacaacaaatacactgcccattgcttcttattt 1153  
 QY 2479 ttgcctggtataaaaaagaacataactctgagtgatgacacagattctgtcgtcctatgaca 2558

Db	1154	ltcgctcgggataaaagaacaaatactctcgtactcttgacaagatcttcgtccctcaatgaca	1213
Oy	2539	aatccgcgaagagatgctgagagctcttgcccttgagagacttccacaagcgcaaacgctgc	2598
Db	1214	aattccgcgaagagatgctgagagctcttgcccttgagagacttccacaagcgcaaacgctgc	1273
Oy	2559	agtgacatctgctacatcaacgctctggagagccgtatgtgtctcagataagccgatctgtcc	2658
Db	1274	agtgacatctgctacatcaacgctctggagagccgtatgtgtctcagataagccgatctgtcc	1333
Oy	2659	tttccatgaaagatgaaagcgagggcgaggtctcatcttgcccttcaaacacacacacattac	2718
Db	1334	tttccatgaaagatgaaagcgagggcgaggtctcatcttgcccttcaaacacacacacattac	1393
Oy	2719	cggcgcgtctgtgagctcccccagagcgcgcgaagcgccgcgtgtggagaaacgcgtgtgagacaag	2778
Db	1394	cggcgcgtctgtgagctcccccagagcgcgcgaagcgccgcgtgtggagaaacgcgtgtgagacaag	1453
Oy	2779	gcaagccagacacatacgaacttctctcaacgcgagacttaacctgatctgcgctctcaacatac	2838
Db	1454	gcaagccagacacatacgaacttctctcaacgcgagacttaacctgatctgcgctctcaacatac	1513
Oy	2839	acccgctctgcgacttctctctaatcccaactctcaacccatgctcaacgtactatactgctca	2898
Db	1514	acccgctctctccatcttctccaaactcccaactcccaacccatgctcaacgtactatactgctca	1573
Oy	2899	tcccaagtatctgcgcctcgatctgtgagagacataatatacagtaataatactgctata	2958
Db	1574	tcccaagtatctgcgcctcgatctgtgagagacataatatacagtaataatactgctata	1633
Oy	2959	tgta 2962	
Db	1634	tgta 1637	

XX	RESULT	3
XX	AAK76375	
ID	AAK76375	standard; cDNA; 2712 BP.
XX		
AC	AAK76375;	
XX		
DT	04-AUG-1999	(first entry)
XX		
DE	Zea mays	SUI starch debranching enzyme encoding cDNA clone.
XX		
KW	Zea mays;	SUI: starch debranching enzyme; maize; endosperm; corn;
KW	bacterial isomylase;	alpha-amylase; starch hydrolytic enzyme;
KW	sugary 1;	ss.
XX		
OS	Zea mays.	
XX		
PH	Key	Location/Qualifiers
ET	CDS	1..2457
ET		/*lag- a
XX		
PN	US5912413-A.	
XX		
PD	15-JUN-1999.	
XX		
PF	24-MAR-1995;	95US-0410784.
XX		
PR	24-MAR-1995;	95US-0410784.
XX		
PA	(IOWA )	UNIV IOWA STATE RES FOUND INC.
XX		
PI	James MG,	Myers AM;
XX		
DR	WPI: 1999-357234/30;	
DR	P-PSDB: AAK17523.	
XX		
PT	SUI starch debranching enzyme	
XX		

PS Claim 3, Fig 1; 39pp; English.  
XX  
XX The present sequence encodes a starch debranching enzyme active in  
CC maize endosperm (Zea mays), designated su1. su1 protein is useful as  
CC a replacement for the bacterial and fungal enzymes currently used in  
CC the starch processing industry.  
XX  
XX Sequence 2712 BP; 636 A; 646 C; 745 G; 685 T; 0 other;

Query Match	33.48;	Score 1000;	DB 20;	Length 2712;
Best Local Similarity	65.78;	Pred. No. 5.4e-206;		
Matches 1883;	Conservative 0;	Mismatches 350;	Indels 633;	Gaps 4

[illegible]

QY	976	aactcgctggcgctgcatcgcccaaaatctgctccaaaactcttgtaagagagctcccaaaagc	1035
Db	1033	gactcgtgactgtctgctcccaaaaatgtaatctaaagctctgtgtaagagagagcccaaaaagc	1092
QY	1036	ggaattgaagctgatacccggaatgtctgtctccaaaccatacagctcgaaagctaatgtagatgt	1095
Db	1093	ggaaattgaggtgatacagagctgtctgtctccatacatacagctcgaaagctaatgtagagaaagc	1152
QY	1096	ccaattatcatcttaaggggctcgataaabaacatactatactatgactgtgaccccaagga	1155
Db	1153	ccaattatcatccttaaggggataagataaatactatacctatacctatgactgtgacccaagga	1212
QY	1156	gaatttataactatctctgctctggaagaactccaaactgtaatacactcggtgttcgt	1215
Db	1213	gaatttataactatctctgctctggaaaaabacctcaattgttaatacactcggtgttcgt	1272
QY	1216	caattcatctgaagctgtcttaagaatcgctgtagcagaaatgcatgtgtgatttcgt	1275
Db	1273	gaatttatagtgattgtctgtgagatacctggtgaacagaaatgcatgtgtgttcgt	1332
QY	1276	cttgactctgcataccaataatgacagaagctccagctctgtggaatccagttaaagctat	1335
Db	1333	cttgactctgcataccaatacagacaagaagatgtagctctgtggaatccagttaatgtat	1392
QY	1336	ggagctccaataagaggttgacatgatacacaagaggaacactctgttaactccacaact	1395
Db	1393	ggaaagcccaatgtagaaggttgacatgatacagaaggaacactctgtgtgccccacact	1452
QY	1396	aatgacatgatacagaatgacaccaatctctgtagagggctgaagaactcatgtctgaagcatg	1455
Db	1453	atgacatgattagcaaatgacaccaatctctgtagaattgtaagaactcatgtctgaagcatg	1512
QY	1456	gattcagaagagccctctctcaagtcaattccactccatgtgaatggttggcttagag	1515
Db	1513	gattcagaagagctcctctctcaagaagagctccactccatgtgaatggttggcttagag	1572
QY	1516	aattggagaatcccggaacattgtgcgttcaattcaatgaaggaacatgtgattctgt	1575
Db	1573	aattggagaatctcgatcgataccgctgctgattcatcaaaagcgaagatgtgattctgt	1632
QY	1576	ggcttctgcgaatctctctgtggaagtccacacttaaccagtgaaagtgtgtgcaatct	1635
Db	1633	ggcttctgcgaatctctctgtggaagtccacacttaaccagtgaaagtgtgtgcaatct	1671
QY	1636	tgtaatgagtgtagatgataatgcacgtttttatataataccaatgatgatataca	1695
Db	1672	-----	1671
QY	1696	tctaaataataaacaatcatagtgatactgcatgtgctgtgaagaagtattgtat	1755
Db	1672	-----	1671
QY	1756	acactagctgcatatataagttttaacaaccaactgtgccaatggaagaaataggtctt	1815
Db	1672	-----	1671
QY	1816	ctagctatcttatttattgttccgttgataataccactgaaaatctccagcatgcat	1875
Db	1672	-----	1671
QY	1876	cttctaggggggagaaagaactatattgtatctgcccccttaagaagacatctcaaat	1935
Db	1672	-----	1671
QY	1936	tcaatagaaagtctctctctctgaagaagaagaacactcatatcttcatctgtgc	1995
Db	1672	-----	1671
QY	1996	taacttagctgatagtatattgttgaagatgaatgccaatctaattgtctggaataattg	2055
Db	1672	-----	1671

QY	2056	atctgctatccacaattctctattggtttctctctgaaataccaacagtaactgttatc	21155
Db	1672	-----	16711
QY	2116	ggcactgcaacttcttatgttaataccagagaggaagaaactcttgccagctacca	21757
Db	1672	-----caagccaggggggagagagccctggccagctaccg	17050
QY	2176	actctgtatctgcacatgatactttaacactctgctctatcttgtaacataataaagaagt	22350
Db	1706	gctctgtatctgcacacagatagatattcaactcgtctgattctgttccacatacaatagcaagt	17655
QY	2236	acaatttccaaaatgvggagagacaacagagatgagagaaaatcaacaactcttgctgaaat	22950
Db	1766	acaacttgcacaaatgvgagagactccagagatgvggaaaaatcataactcttgctgaaat	18350
QY	2256	gtgvggagagaaagagaattccgcaagattgtctctcaaaaagatctgvggaaagvgccagatgc	23555
Db	1836	gtgvggagagaaagagaaattctgcacaagtctctgcacccgaaagataagagaaagccaaatgvc	18850
QY	2356	gcaatttcttggattctgtctcatatggtttccaaagaaatttccaaatgtcttccaaatgvggtgtg	24150
Db	1886	gcaatttcttctgtctgtctctatagttctctcaagagagttccaaatgtcttccaaatgvggcgtgc	19450
QY	2416	aataatgcccacaacaaaagvggggcacacaacatactacatcgtgcattgtatctgtcaaat	24757
Db	1946	aataatgtcacacaacaaagvgggagagacaacatacgtatcgtccatgvcacatattgtcaaat	20050
QY	2476	atttctgctggtgataaaaa--aagaaacaatactcagatctgcacagatctcgtctccaa	25330
Db	2006	attctcgtctggtgataaagaaagaaagaaacaaactcctgattctgtacagatattctgcgtccaa	20650
QY	2533	tgcacaaaattcccgaaaagagatgvgcagvggtctctgaccttgaaagaaattcccaaaagcccaaa	25950
Db	2066	tgcacaaaattcccgaaaagaaatgvgaaatcctctcgtgcctctgcagagactcccgaaactccaagaac	21250
QY	2593	ggctgcgaatgvcacatgctcatcagcctctgcgaaagccctgattgtgtctcgaaataagcccgatctcg	26520
Db	2126	ggttgtaaaatgvcacagatcatcagcccggtgagagctcgtacacgtgtcagaaagvgaaagccgatactcg	21850
QY	2653	tgtgccttccatgaaaagatgaaaagaaagagcagagatactatgtgaccttcaaacacagacc	27120
Db	2186	tgtgccttccatcacaatgaaagaaagaaacaaaagcagagatactacagatggtgccttcaaacacacagtc	22450
QY	2713	acttaccggtccgtgtctgtgagatctcccaagagccgvcgaagccgctgcgggtgvggaaacgggtgtgvgg	27772
Db	2246	aaccttcgggtgtgtgtgtgtgtgtcttccaaagagcgtcctgtgtctcggatgvggagccgggtgtgtgtg	23050
QY	2773	acaacagvgaaagccagacacacatacgaacttcttccacgaaacgaacttaacctgtatctgcgtctcca	28320
Db	2306	acaacagvgaaagagagccacacataatgaaacttccctccacgaaatggtcctgcacaaagatctgtgtctcca	23650
QY	2833	caatacacaagaattctgcgaatttcttctctacacacaaaccttccaaacacatgcttccagctatcatat	28920
Db	2366	ccgtctacacagttctctcatattctctcaactccaaactctctatctatcctatgctccagctatctct	24250
QY	2893	cggtcatctctagatctgcgccctgagatgtttgagagacaataata	2938
Db	2426	ccatactctgtattctgcgccctgtatctctgaaagaaagagaaata	2471

RESULT 4  
AAx27061  
ID AAX27061 standard; DNA; 2706 BP.  
XX  
XX  
AC AAX27061;  
XX  
XX  
DT 21-MAY-1999 (first entry)  
XX  
XX  
DE S. tuberosum isoamylase clone I5 coding sequence.  
XX  
XX  
XX Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;  
KW starch modification; debranching activity; ss.

XX		Solanum tuberosum.	
OS		WO9906575-A1.	
XX			
PN		11-FEB-1999.	
XX			
PD		30-JUL-1998;	98MO-GB02280.
XX			
PX		31-JUL-1997;	97GB-0016185.
XX		(PLAN-) PLANT BIOSCIENCE LTD.	
PA		Bustos Guillen R, Edwards EA, Martin CR, Smith AM;	
PI		WPI; 1999-153803/13.	
DR		P-PsDB; AAY00869.	
XX		New nucleic acid from potato encoding starch debranching enzyme with	
PT		isoamylase activity - and related vectors, transformed cells and	
PT		plants, proteins and antibodies, used to generate starch with	
PT		altered properties, for use in foods, as thickeners etc.	
PS		Claim 3; Fig 2; 78pp: English.	
XX		This sequence encodes a Solanum tuberosum isoamylase of the invention.	
CC		Fragments of the isoamylase coding sequence are used to identify and	
CC		clone isoamylases from other plant species, by standard hybridisation	
CC		or amplification methods. Expression of the isoamylase DNA in host cells	
CC		is used to produce the isoamylase which is used: (1) to raise antibodies	
CC		(Ab) for identification, isolation and localisation of isoamylases); and	
CC		((1) for synthesis of branched polysaccharides. The DNA, proteins	
CC		and Ab (or related peptides) are used to alter the quality and	
CC		quantity of polysaccharides in a host cells, specifically to alter	
CC		branching in amylopectin. The modified starches produced are useful in	
CC		human or animal foods; as biodegradable plastic; as food or paint	
CC		thickener; in starch-coated films, paper and textiles; in mixing	
CC		explosives; in pharmaceuticals and glues. Manipulation of debranching	
CC		activity in a plant allows control of starch properties, e.g. increased	
CC		gel strength; formation of paste rather than gel, changed physical	
CC		characteristics etc.	
XX		Sequence 2706 BP; 797 A; 492 C; 624 G; 792 T; 1 other:	
SO			
	Query Match	22.5%; Score 673.2; DB 20; Length 2706;	
	Best Local Similarity	69.4%; Pred. 1.8e-135;	
	Matches 937; Conservative	0; Mismatches 398; Indels 15; Gaps	1
OY	395 aggataaggttgcacggagggtttcccttcgacccttgatgaactcgacccggaaagtgt	454	
Db	425 aaagaaggattgacgcggagcaatttctcgtgatctcttagctaataactyggagattgt	484	
OY	455 ggcattcttccttcgaaagcagctgcacaacaatgtttagggtaacagttcgacagcca	514	
Db	483 ggcatgttgttctcctaagsgaatgttgagaatatgtctatagctacaatlttgatggga	544	
OY	515 ccttgcgtcccaactcggygcactacctgttatcttccaatgttcgtgtgagatccttatg	574	
Db	545 aattctgtctgaagaagaaacactattcttgactcttcgcagatagtggttgatccttatg	604	
OY	575 cttaaggcagttgataagccgaagggaattatggtgtccacgcgcgtgtaacaattgtctgc	634	
Db	605 ccaaaggctatgaataagcagaagggaataatggtatattlaaggccaaaggagatattgttgc	664	
OY	635 ctcaagaatctgcagatgcccccctctccatatagcacagctttgattcggaaaggacactac	694	
Db	665 ccccaaagtgcgtgcatgttaccttctgcgttccttgatcagtttgttggaagsgactac	724	
OY	695 ccttaagatatccataaaaagacgtgtatataatagatagcaacttgcgttgatccaaga	754	
Db	725 caactgaagttccacagagagaatctctgtatattcataaataatgcttgcgtgggttacta	784	

QY	755	agcagatctcaagcaacgctagaaacatccgggtaccttccatctggaagctgtgcgcgaagctgtg	814
Db	785	atcagagctcgagtggaacaacaaatattccgtgtacttcaactctgtgtcttgtaggaataactgtg	844
QY	815	actattgaagagacttgtagaattatgttatgttaattgaatggccctccatctggaatcaag	874
Db	845	atcacttgaagaaactctgtgtgtccaaactgttataaagcttcaatgccctgtccagagttcaag	904
QY	875	agcttgaagctaccacaacctctctctc-----caagatgaataactcttgtagaatt	919
Db	905	agcttgaagctactatagttatataactctgttatcttggtgcgaactcaagaagtttaactcttgggct	964
QY	920	attctaccataaactctcttctccacaaatgaacaagatacacatacgaagcggaataaaaaact	979
Db	965	attctaccatgaattctcttctctccaaatggaagataactcatctgtgtcttaagtaatt	1024
QY	980	gtggagctgtagatgcacataatgagttcaaaactttgttaagagtggtctccaaacgggga	1033
Db	1025	ggcgccctcggtagcaataaagcaatttaagttatctgttcaaggagagacataaaacgtgga	1084
QY	1040	ttgaggtgacccttgtagtgtgtcttccacaacatacagcttgagggttaaigtgaatgttccaa	1099
Db	1085	tcgagagttatcaatgtgagtggttttccaaatcaacactgcttgaaagaaatgaaatgttccaa	1144
QY	1100	tattatcatttaaggggggtcgaataatactacatactatagctgttcgacccaaggggaagt	1155
Db	1145	tactatcatcttaaggagcatctgacaaacagtggtgtttatatacgtatgcttccataaggttgat	1204
QY	1160	tttataactatctcgtgctgtgtgggaatactcttcaactgttaatactcctgtgtctgtcaat	1211
Db	1205	tttaacaacatactcagagatgtgtggaatactcttcaactcgttaatactccatctgtacat	1264
QY	1220	tcaatgtgagatgttttaagaatacagtggtgacgggaataatgcagtgtatgtgtttctgtttg	1279
Db	1265	ttatagtgagatcttgcttgaagataatctgtgtatccgaaatgcagctgaatgtgtctccgttgg	1322
QY	1280	atctctgcataccataatgaaccagaggttccagactctgtgtgggaatccagatgaacgtgtatgaag	1335
Db	1325	atctgtcttctatccttcaacaagaagtacaggtcgcgtgtgaagtgtgttaaatgtctaatgga	1384
QY	1340	cttccaaatgaaggtgtgacatgtatccaacaagaggaacacctctgttactccacaactatgtg	1399
Db	1385	attcaattgacggtgtgacgtgtatccacaagacactcctctccacaagcccaacatctgattg	1444
QY	1400	acattgatacgaatgagccaactttcttgtagggcttcaagctcattgtttgtagaagttgattg	1455
Db	1445	ataatgattagaagaagatacccaatactctcgtgagtaaaagcttcaataagcttgaagaatcgtggaat	1504
QY	1460	cagggagcgtctatcaagaatgagttcaactctccctacacgtgaagtgtttgtctgtgtgaagt	1519
Db	1505	gtgtggagccttaccacaagttgtgcagtgttcttcgcaacgtgggttatctgtgtccgggtgtgaacg	1566
QY	1520	ggaagtacgggaacatctgtgcgttcaattcaataaagagcacatgtgaattgtgtctgtgtgct	1579
Db	1565	gaagaatcaggatcagatgtatccgagattcaatccaagaagcacatgtgtgtttcttgcgggctt	1624
QY	1580	ttgcgcgaatgtccttgttgtagaagtccacaactatacacaagttaagtgtgtgcaatattgtta	1633
Db	1625	ttgtctggaatgtccttctgttggaagcccaaatctatacacaagaagaagtagaagaaaaacaattga	1684
QY	1640	aatgtgtgagtgaaatgtccacctgtgaatttttataataccacaatgatatatacatata	1699
Db	1685	acagatataaattctcgtgtgtgtgcaccaagatgtgttctccttctgtgcgtatattagtacataca	1744
QY	1700	ataataataacatcatagatgtatgcatatg	1729
Db	1745	acaataaacacaattgtgcaaatgtgagagg	1774
RESULT 5			
AAAX27066			
ID AAAX27066 standard; DNA: 2806 BP.			

AC AAX27066;  
 XX  
 DT 21-MAY-1999 (first entry)  
 XX  
 DE Original S. tuberosum isoamylase clone 15 coding sequence.  
 XX  
 KM Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;  
 KM starch modification; debranching activity; ss.  
 XX  
 OS Solanum tuberosum.  
 XX  
 PN W09906575-A1.  
 XX  
 PD 11-FEB-1999.  
 XX  
 PF 30-JUL-1998; 98WO-GB02280.  
 XX  
 PR 31-JUL-1997; 97GB-0016185.  
 XX  
 PA (PLAN-) PLANT BIOSCIENCE LTD.  
 XX  
 PI BUSTOS GUILLEN R, EDWARDS EA, MARTIN CR, SMITH AM;  
 XX  
 DR WPI: 1999-153803/13.  
 XX  
 DR P-PSDB: AAY0872.  
 XX  
 PT New nucleic acid from potato encoding starch debranching enzyme with  
 PT isoamylase activity - and related vectors, transformed cells and  
 PT plants, proteins and antibodies, used to generate starch with  
 PT altered properties, for use in foods, as thickeners etc.  
 XX  
 PS Disclosure; Fig 11: 78bp; English.  
 XX  
 CC This sequence represents the original isolated clone encoding a  
 CC Solanum tuberosum isoamylase of the invention.  
 CC Fragments of the isoamylase coding sequence are used to identify and  
 CC clone isoamylases from other plant species, by standard hybridisation  
 CC or amplification methods. Expression of the isoamylase DNA in host cells  
 CC is used to produce the isoamylase which is used: (1) to raise antibodies  
 CC (Ab) (for identification, isolation and localisation of isoamylases); and  
 CC (11) for synthesis of branched polysaccharides. The DNA, proteins  
 CC and Ab (or related peptides) are used to alter the quality and  
 CC quantity of polysaccharides in a host cells, specifically to alter  
 CC branching in amylopectin. The modified starches produced are useful in  
 CC human or animal foods; as biodegradable plastics; as food or paint  
 CC thickener; in starch-coated films, paper and textiles; in mining  
 CC explosives; in pharmaceuticals and glues. Manipulation of debranching  
 CC activity in a plant allows control of starch properties, e.g. increased  
 CC gel strength; formation of paste rather than gel, changed physical  
 CC characteristics etc.  
 CC  
 XX  
 SO Sequence 2806 BP; 818 A; 523 C; 654 G; 807 T; 4 other;

Query Match 20.5%; Score 613.2; DB 20; Length 2806;  
 Best Local Similarity 68.0%; Pred. No. 1,6e-122;  
 Matches 922; Conservative 0; Mismatches 410; Indels 24; Gaps 4;

QY 395 agaatgggagggacggaggggtcccttgacccctgagatgaacggagcggagcgt 454  
 DB 488 agaagagggagggacggaggggtcccttgacccctgagatgaacggagcggagcgt 547  
 QY 455 ggaatgctcatcgaagggagggcgcacaaatcttcaaggttcaaggttcgacggca 514  
 DB 548 ggaatgctcatcgaagggagggatttgagaataatgctatagtctacaatcttgagg 607  
 QY 515 ccttgctcctcactcggcgacgaactctgaatgttccaatgtcgtgtggaatccctatg 574  
 DB 608 aatctcctcgaaggaaggaactactctgacactcttcgacagatagtttgatccctatg 667  
 QY 575 ctaaggaaggaatgaagcggagggagatagtggttccagcggcggtgaacaatgctggcg 634  
 DB 668 ccaaggtatagtaagcagaaggaataatggttatattaggccagaggaatgattctgtgc 727

QY 635 ctcaagatgctgagcagatccctct---tccatatgacacgtttgttggaaaggcgacc 691  
 DB 728 ccccaatgctgagcagatggtacccctctctgagatcagttgattggaaggagatc 787  
 QY 692 taccatgaatatccca---aaaggacctgttatatgtgagatgacatctgagat 748  
 DB 788 taccacatggaagcttcccaagaagaatctctgtatcmaatgaatgaatgacatgctcgatg 847  
 QY 749 tcaagaagcagatgtcaagaacatgtagaacaatccgggtacttcatatgagctgtgcga 808  
 DB 848 ttactatccatgtctcgagtgaaacaaataatctgttacttactctggtgtgtgga 907  
 QY 809 agcttgactatcttgaaggagcttggaatattgtaattgaatcaatgacctgcacatg 868  
 DB 908 aacttgactcttgaaggagactgtgtcacaactgtatagatcaatgacctgtgcacagat 967  
 QY 869 tcaagagctggaggtactcaactctctc-----caagatgaacttt 913  
 DB 968 tcaatgagctggaggtactcaactctctc-----caagatgaacttt 1027  
 QY 914 ggggatatctacacataaactctcttccacaatgacaagaatcacatcagcggtgataa 973  
 DB 1028 ggggctattctactgtcaattctcttctccaatggagaagatactcctcgtgtctaa 1087  
 QY 974 aaactgtggcgatgagcagataaagatgacaaactttgtgaaggagctcacaac 1033  
 DB 1088 gtaattcgccctcggtgacataaagcaatttaagtaatctgtcgaaggagacataaac 1147  
 QY 1034 ggggaattggaggtatctctgtgtgtgttccaacacacagcgaggtgaatgagat 1093  
 DB 1148 gtggaatcgaaggtatcaatgattgttcttccacacacgctcgaggaagaaatgaaatg 1207  
 QY 1094 gtccaatattacatttaagggggtccgataataacataactaatatgtctgcacccaag 1153  
 DB 1208 gtcccatattacatttaagggggtcattgacaaacgtgtgtttaaagctgtccttaag 1267  
 QY 1154 gagaatttataactatctctgtgtgtggaataccttcaactgaatcctgtgttc 1213  
 DB 1268 gtgaatttcaactacactcagatgtggaataactcctcaactgaatcctgtgttc 1377  
 QY 1214 gtaattcatgtatgattgtttaagatactgggtggaaggaatgacgtgtgtgttc 1273  
 DB 1328 gtcaat---ttaaagtgtcgtgagatattggtgtaccgaaatgacgtgtgttc 1384  
 QY 1274 gtttgatctgtgcacataatgacagaggttcacgtctgtgtgacccaatgacgt 1333  
 DB 1385 gcttgcattctgtctcactcactcacaagaagtagcagctgtgaaatgctgtatgtct 1444  
 QY 1334 atggaatcacaatagaaggtgacatgatacacaagaagacacactctgttactcaacac 1393  
 DB 1445 atggaatcacaatgacgtgagtgatgacacacagacacactcctcacaagccacacat 1504  
 QY 1394 ttattgcagatcagaacatgacccaattcttggagcggtcaagctatgtctgaagcat 1453  
 DB 1505 tgaattgatattagaaatgataatcactctgtgagataaactctatagctgaagcat 1564  
 QY 1454 ggaatgcaagagggcctcatcaagtagtcaatccctcactcactggaatgtgtgtgacgt 1513  
 DB 1565 ggaatgttgagagggccttaccaaagtgtgacatgttccgcactgtgtgtatcgtgcgagat 1624  
 QY 1514 ggaatggaaggtacgggacatgtgtgcataatcattaaaggacatgattgtgtcgt 1573  
 DB 1625 ggaacggaagaatgacggagacatgtgtgacagcattcacaagaagcagatgtgttcctg 1684  
 QY 1574 gtgttttgcggaatgtccttgttggaaggtccacacccataaccaggtgaatgtgtgcata 1633  
 DB 1685 gggcctttgtcgaatgcttgttggaagcccaactctactaccagaagggaggaagaaac 1744  
 QY 1634 ctgttaatgattgtgaatgacgtcacctgtgattttatataacacatgatatca 1693  
 DB 1745 catggaacagatataaattcgtgtgtgcacagatgtgttacttggctgtatttagtga 1804





XX	AA34655;
AC	
XX	
DT	05-JUL-1999 (first entry)
XX	
DE	Hexaploid wheat DBE genomic sequence.
XX	
KW	Starch biosynthetic pathway; cereal plant; enzyme; SBE, SSS, DBE, GBSS;
KW	starch branching enzyme; starch soluble synthase; debranching enzyme;
KW	endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
KW	grain softness protein I; bacterial isoamylase; glycogen synthase;
XX	WSBE I-D4 gene; ss.
OS	Triticum tauschii.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..9287
FT	/tag= a
FT	/note= "the start codon is not indicated; contains
FT	146 internal stop codons"
PN	W09914314-A1.
PB	25-MAR-1999.
XX	
PF	11-SEP-1998; 98WO-AU00743.
XX	
PR	20-MAR-1998; 98AU-0002509.
XX	
PR	12-SEP-1997; 97AU-0009108.
XX	
PA	(CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA	(GOOD-) GOODMAN FIELDER LTD.
PA	(LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PA	(AUSU ) UNIV AUSTRALIAN NAT.
XX	
PI	Li Z, Morell M, Rahman S;
DR	WPI; 1999-229525/19.
DR	P-PSDB; AAY06919.
XX	
PT	New isolated cereal plant enzyme genes used for, e.g. expression of
PT	antisense sequences of granule bound synthase
XX	
PS	Clalm 14; Page 103-115; 171pp; English.
XX	
CC	The invention relates to a novel enzyme of starch biosynthetic pathway
CC	in a cereal plant, where the enzyme is selected from starch branching
CC	enzyme (SBE) I , SBE II, starch soluble synthase (SSS) I, and debranching
CC	enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
CC	SBE I of rice or maize. The methods and products can be used for
CC	targeting expression specifically to the endosperm of the seeds of cereal
CC	plants such as wheat or barley. They can be used for the expression of
CC	e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
CC	moli. wt. glutenin, grain softness protein I, bacterial isoamylase,
CC	bacterial glycogen synthase, and wheat high moli. wt. glutenin Bx17. They
CC	can be used for modifying the characteristics of starch produced by a
CC	plant. The present sequence represents the wheat DBE genomic DNA
CC	sequence.
XX	
SO	Sequence 9289 BP; 2477 A; 1909 C; 2040 G; 2863 T; 0 other:
	Query Match 17.8%; Score 533.4; DB 20; Length 9289;
	Best Local Similarity 95.1%; Pred. No. 3.9e-105;
	Matches 562; Conservative 0; Mismatches 26; Indels 3; Gaps 1
Oy	1 ggtcggggcccgcgcgcctcgcagcgttgccgaacccaatcgacgcggggaaaggggg 60 
Db	1943 ggccggggggcgcgcgcgcctcgcagcggttgsgaaccaatgacgcggggaaggggg 2002
Oy	61 tcggcgagagtgtgcgcgcgcggttcttcgaagcgcgacgagaagtataaggaagag---gggg 117 
Dh	2003 tctggcgagaggtgtgcgcgcgcggttcttcgaagcgcgacgagaagcgcgagatatgaaggaagcgg 2062

QY	118	agcgagacgagacgcgtctgcgagagacaggtacgcgcctcgcgcgcgtctgaaagtctcgc	177
Db	2063	acgagagagagaaagcggctggcggagagacaggtacgcgcctcgcgcgcgcgtgcaaggctcgc	2122
QY	178	ccggaatgcgccgcgcctcgcgcgcacgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcct	237
Db	2123	ccggaatgcgccgcgcctcgcgcgcacgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcct	2182
QY	238	atccgc	297
Db	2183	atccgc	2242
QY	298	tggcggttcgcctcccggaatagaatcacaacgttttgctgtgcgcgcgcgcgcgcgccttttgc	357
Db	2243	tggcggttcgcctcccggaatagaatcacaacgttttgctgtgcgcgcgcgcgcgcgccttttgc	2302
QY	358	gcccgcgaatttaagctttctgtactcgtgggcgaacgtctgcgaagataagggtgtgacgcgagagctt	417
Db	2303	gcccgcgaatttaagctttctgtactcgtgggcgaacgtctgcgaagataagggtgtgacgcgagagctt	2362
QY	418	cccccttgacccccctgatgatcgcgaaccgcgggaacgtgtgcacatctctcatalcgaagcgag	477
Db	2363	cccccttgacccccctgatgatcgcgaaccgcgggaacgtgtgcacatctctcatalcgaagcgag	2422
QY	478	ctgcacaacaacagcttttcgcgttaacaggttcgcgaagcacccttgcctcctacatcgcgggac	537
Db	2423	ctgcacaacaacagcttttcgcgttaacaggttcgcgaagcacccttgcctcctacatcgcgggac	2482
QY	538	taccttgatctgttccaatgtctgtgtgataccttatcgtacaagcaatgata	588
Db	2483	taccttgatctgttccaatgtctgtgtgataccttatcgtacaagcaatgata	2533
RESULT 8			
AAK33162			
ID	AAK33162	standard; DNA: 1663 BP.	
XX	AAK33162;		
XX	06-JUL-1999	(first entry)	
DT	Potato isoamylase type DBE genomic clone SEQ ID NO:14.		
XX	Potato: isoamylase debranching enzyme; DBE: starch; ss.		
DE	Solanum tuberosum.		
XX	WO9912950-A2.		
OS	18-MAR-1999.		
XX	04-SEP-1998;	98WO-GB02665.	
XX	06-SEP-1997;	97GB-0018663.	
PR	(NMTT ) NAT STARCH & CHEM INVESTMENT HOLDING COR.		
XX	Jobling SA, Schwall GP, Westcott RJ;		
PI	WI: 1999-229220/19.		
XX	P-PSDB; AAY04382.		
DR	New isolated potato isoamylase-type debranching enzyme gene		
XX	Claim 7; Fig 8; 72pp; English.		
PS	The present invention describes a novel nucleic acid sequence which is		
CC	obtainable from potato plants and carries at least a portion of an		
CC	isoamylase-type debranching enzyme (DBE) gene. Constructs containing the		
CC	nucleic acid sequence can be used to alter the starch properties of		
CC	plants such as potato, sweet potato, maize, wheat, barley, oat, cassava,		
CC	pea or rice. The starch can have increased branching and/or shorter		



```

OY 649 atgatacccttccataatgacagcttgatgggaagcgacactctctaagatctct 708
DB 674 ttaataagcttgctcttcttgctg-----gagaaataatcaagcttccaatatacc 727
OY 709 caaagaagaccttgtaataatgagatgacttgctgtagatccaagaacatgatacagc 768
DB 728 gagaagaagactcttataatctgagatgaaatgcttgcgtcttactgctgatacacaagc 767
OY 769 a-----atgtagaacatcccggttacttcatatggagctgctgtaagctgtaactatg 822
DB 788 agcttgatcaagaatcaacgggaaggttaccttgcttaattgaaagataccacatctt 847
OY 823 aaggaacttggaatgtaattatgtaattatgcccctgcatggtgtccaagaactggag 882
DB 848 ctggaacttggtgtaactgagtagaattatgctctgttttggagcttgatgaactggaa 907
OY 883 tactcaacacctctctc-----caagaatgaacttctgggagatattccacata 930
DB 908 ttcaaaaggagacctaataccgagagatcacatgatacatatgaggctactcaacata 967
OY 931 aactctcttcaacaatgacaaatatacacatcaggcgagataaaaaactgtggcgatg 990
DB 968 aactcttctgccaatgagctgcatgacagatgtg-----tgcgagccgtccgt 1021
OY 991 gccataatgagttcaaaacttctgtgaagagagctcaacaacgggaattgagtgatc 1050
DB 1022 gcttcctggaggttcaagaatggttcaaggcttgatggtgctggaattgaaagctc 1081
OY 1051 ctgagatgttcttcaacacacacagctgaggtgaatgaaatgtcca---atatcata 1107
DB 1082 ttgagatgtcttataatcacacaatgaagctgagtgatgaataaccatacactcata 1141
OY 1108 tttaaaagggttcgataatactatactatgctg-----accacaaggagagctt 1161
DB 1142 ttccgggaatagacaacaaggtttatctacatgtagatttaaacacaacatgctcagctg 1201
OY 1162 tataactatctgctgctggaataccttcaactgtaatacctctggtcgtcaatc 1221
DB 1202 ctgaatctcgtgagtgatgaataacttcttaactgcaacacccacaagctatggaact 1261
OY 1222 atgtgagatctgtaagaatctggtgaagaaatgcaatgctgtagtcttcgcttgat 1281
DB 1262 atacttgaagaactgaagacacgtgggtccacagatcatactgctgagatctcgcttgat 1321
OY 1282 ctgtgacatcaataatgacagaggttccagctgctggaatccagatlaacgtgtatgagct 1341
DB 1322 ctgtcagatgttctcttctgacagag----- 1344
OY 1342 ccaatagaaggtgacatgatacaacaaggaacacctgttactccaacactaatgac 1401
DB 1345 -----gacagatgttactcccatlaatgctctccctctgttaag 1384
OY 1402 atgatacgaatgacccaatctctggagcgctcaagctcatgcttgaagcattgagatga 1461
DB 1385 gccatttcccaagaatgtagtctatgttcgaagtgcaaaatattctgtgagccatggagatgct 1444
OY 1462 ggaagcctctataatgagtagtcaatctccctcactggaatgttctgctgagtggaatgg 1521
DB 1445 ggaagcctctataatctctgtggaagttcccgaaactgggaacgggtgctgagtggaatgg 1504
OY 1522 aagtaacgggaacatgtgctcaatctcaatlaaaggaacatgtagatctgctggtggttct 1581
DB 1505 aagtaacgggaatgacatcaagaagaaattataaaggtcgatgctgacatgaagaatctt 1564
OY 1582 gccgaatgctcttggaagtcacacactatacagagtaa 1621
DB 1565 gcaacccgtatcgaggttccagcgatctgtacagagtga 1604

```

RESULT 10  
 AAC81914/c  
 ID AAC81914 standard; DNA: 273254 BP.  
 XX

```

AC AAC81914;
XX
DT 27-FEB-2001 (first entry)
XX
DE Chlamydia pneumoniae genome DNA.
XX
KW Genome; diagnosis; vaccine; ds.
XX
OS Chlamydia pneumoniae.
XX
PN MO200027994-A2.
XX
PD 18-MAR-2000.
XX
PE 12-NOV-1999; 99WO-US26923.
XX
PR 12-NOV-1998; 98US-0108279.
PR 08-APR-1999; 99US-0128606.
XX
PA (RECC ) UNIV CALIFORNIA.
XX
PI Stephens R, Mitchell W, Kaiman S, Davis R;
XX
DR WPI: 2000-376516/32.
XX
PT Isolated nucleic acid for use in diagnostic and analytical methods
XX
PT encodes genomic sequence of Chlamydia pneumoniae -
XX
PS Claim 2; Page 128-320; 320pp; English.
XX
CC This invention describes a novel nucleic acid (N1) encoding a Chlamydia
CC pneumoniae protein (P1), given in the specification. The isolated nucleic
CC acid is useful for diagnostic and analytical methods, such as,
CC hybridization-based assays or amplification-based assays. The protein may
CC be used for diagnostic purposes, for their enzymatic or structural
CC activity, or as a vaccine. The invention also describes (1) a probe
CC comprising a hybridizing fragment of N1; (2) an isolated nucleic acid
CC (N2) that hybridizes under stringent conditions to N1; (3) an expression
CC cassette comprising N1 under the transcriptional regulation of a
CC transcriptional termination region; (4) a cell comprising an expression
CC cassette of (3) as part of an extrachromosomal element or integrated into
CC the genome of a host cell as a result of induction of the expression
CC cassette into the host cell, and the cellular progeny of the host cell;
CC (5) a method for producing a P1 comprising growing a cell of (4) where
CC the protein is expressed and isolating the protein free of other
CC proteins; (6) a purified polypeptide composition comprising at least 50
CC weight % of P1; and (7) a monoclonal antibody binding specifically to the
CC peptide of (6).
XX
SQ Sequence 273254 BP; 76423 A; 51054 C; 61965 G; 83812 T; 0 other;
XX

```

Query Match 4.8%; Score 143; DB 21; Length 273254;  
 Best Local Similarity 55.3%; Pred. No. 2.2e-20;  
 Matches 359; Conservative 0; Mismatches 260; Indels 30; Gaps 3;

```

OY 671 cgttgaatgggaagcgacactacccttaagaatccctcaaaaggagactgtaataatg 730
DB 170596 CATTTCCTTGGGAGGATGATCAGCCTCTGCATTTGCCGAAGAAGATGATCATCTAGT 170537
OY 731 agatgacactgctggatctcaagaagcatgattcaagaacatgtaacatccggatct 790
DB 170536 AGATGCATGATCGTTCCTTCACGCAATCTCTCATCTAGAGGTTCATCTCCGCGAACC 170477
OY 791 tcaattggagcgtgtcgaagcttgactatgtgaaggaactggaattatgattgagat 850
DB 170476 TCTTAGGAATCATTTGAAGAATCGACCATCTGCATTAACCTGGGAATCAACCGTGTGAAC 170417
OY 851 taatgcctgcacatgagttcaacagagctggag-----tactcaacacctctct 898
DB 170416 TCTTAGCATCTTTGAGTTCATGATGAGACTGGCATCTTTTAGAATTCGAATTCCTCTT 170357

```

Qy 899 ccaagaagaacttttgggagatattctacataaacttttccaccatgacagaatata 958  
 Db 170356 ATCTGTCAATTAATGGGGTATGCTCCCTAAATTTCTTCTTCCGACGTTATG 170297  
 Qy 959 catcagcgagataaaacactgtgscgtgagtcacataaagttcaaacactttgtcaa 1018  
 Db 170296 CTTATGCGC-----TCTATCTTCTGCGCTCCCAAGTAGAGAGTTTAAACTTTAGTAA 170246  
 Qy 1019 gagaaggtcacaaacgggggaaattgagtgatcccttgagctgtcttccaacatacagctg 1078  
 Db 170245 AGACCTTGCAATCAAGAGGTATGTAGGTCAATCTTGATGTTGTTTATCATACGGGCT 170186  
 Qy 1079 agggtaattgagatggtcccaattatcattaaaggggtgagataactatactata 1138  
 Db 170185 TGC-----AAGCGACGACCTGCTCTTGGCTTGTGATGACACTCCGACCTATTATA 170135  
 Qy 1139 tgcttgaccccaaggagagatttaactattctgctgtggtggagaaacctcaactgta 1198  
 Db 170134 TTTTAGATGCACAGGTCACTTTACAAATTTATTCAGGCTGTGGAACACTCTCAATPACA 170075  
 Qy 1199 atcattcctgtgtcgtcaattcatctgtatgttttaagatactgggtgacggaaatgc 1258  
 Db 170074 ACCCGCGCCCCACGACCAATGATTCGACATCTTACGTTATGGGTAAGAAAGATGC 170015  
 Qy 1259 atgttgatgttcttcgtttgtatcttgcatacctaatagaaccagaggttc 1307  
 Db 170014 ATGTGATGGGTCCGATTGATCTTGTCTGTCTTCTTCGTCGTC 169966

## RESULT 11

AAx91990  
 ID AAX91990 standard; DNA; 1230025 BP.

XX AAX91990;

DT 13-SEP-1999 (first entry)

XX Nucleotide sequence of the complete genome of Chlamydia pneumoniae.

XX Respiratory disease: pneumonia; bronchitis; heart disease; sarcoidosis;

KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

KW vaccine; neutralising epitope; ss.

XX Chlamydia pneumoniae.

XX WO9927105-A2.

PD 03-JUN-1999.

XX 20-NOV-1998: 98WO-IB01890.

XX 04-NOV-1998: 98US-0107078.

PR 21-NOV-1997: 97FR-0014673.

XX (GEST ) GENSET.

PI Griffiths R;

DR WPI: 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

XX Claim 1: Page 291-611; 1912pp; English.

XX The present sequence represents the complete genome of Chlamydia pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in immunogenic compositions as vaccines. CC Vectors containing C. pneumoniae nucleotide sequences can also be used as immunogenic compositions, especially where the vector directs

CC the expression of a neutralising epitope of C. pneumoniae.

XX Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other:

Query Match 4.8%; Score 143; DB 20; Length 1230025;  
 Best Local Similarity 55.3%; Pred. No. 3.6e-20;  
 Matches 359; Conservative 0; Mismatches 260; Indels 30; Gaps 3;

Qy 671 cgttgatgggaagcgagccctacacttaagataatctcaaaagagcctgtaataatg 720

Db 446195 catttccttgaggatggtgatacgcctctctgcatcttcggaagaagagatcatctaa 446254

Qy 721 agatgcaacttgctggtatcacgaagcatgattcaagcaagaaatgaaacatccggtaact 790

Db 446285 agatgcatgacgtcccttcacgacaaatctctcatcattaggttcatgctccggaaact 446314

Qy 791 tcaatggagctgtgctcgaaagcttgactattgaagaagcttggaattatgtatgaat 850

Db 446315 tcttgagaaatcatgaaagatcgacccaatctgcatagaactggaatcaacgcgtgtgaac 446374

Qy 851 taatgcccgtgcattgattcaacagcgtgag-----tactcaacctctct 898

Db 446375 tcttaacctatcttttgaattcgtatgagctgcgacatccctttgaaatttcgaaattccct 446434

Qy 899 ccaagaagaacttttgggagatattctacataaacttttccaccatgacagaatata 958

Db 446435 atctgtgcaattatttggggtatgctcccccataattctctctctccttgccgagctatg 446494

Qy 959 catcagcgagataaaacactgtggtggtggtggtggtggtggtggtggtggtggtggt 1018

Db 446495 ctatgccc-----tcgtatcctctgctgcctccaaagtgaagatttaaaactttgtaa 446545

Qy 1019 gagaaggtcacaaacgggggaaattgagtgatgagctgagtggtgtctcaacacatacagctg 1078

Db 446546 agacctgcatcaaaaggtatctgaggtcattcctgagtggttttctaatacacaaggtc 446605

Qy 1079 aggttaattgagaattgtcccaatatattcaattaaagggtggtcgaataatacatactata 1138

Db 446606 tgc-----aagggaagacctgtctcttgccttggtgtagaactccgaggtattata 446656

Qy 1139 tgcttgacccaagggagagatttataactattcgtgctgtgtggaaatacctcaactgta 1198

Db 446657 tttagatgacaaaggtccttacaataattacagctgctgtgaacacatctcaatacaa 446716

Qy 1199 atcattcctgtgtgtcgtcaattcatctgtagattgtttaagatactggtggaacgaatgc 1258

Db 446717 accggtcccccaagccaatgattcgtgacatcttaccgttatttgggaagaagaatgc 446776

Qy 1259 atgttgatgttcttcgttttgatcttgcaataaagacccaagaggttc 1307

Db 446777 atgtgcatggtgtccgatttgatcttgcctctctctctctctctctctctctctct 446825

## RESULT 12

AAx27067  
 ID AAX27067 standard; DNA; 1389 BP.

XX AAX27067;

DT 21-MAY-1999 (first entry)

XX Original S. tuberosum isoamylase clone 9 coding sequence.

XX Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;

KW starch modification; debranching activity; ss.

OS Solanum tuberosum.

XX WO9906575-A1.

XX 11-FEB-1999.













```
QY 2116 ggcactgcaactctctatgtatgttaactcagcagaggaagaaacctggccagatca 2175
Db 1672 -----CAGCGAGGGGAGGCGCTTGCGACAGTATCG 1705
QY 2176 acctgtatgtgcacatgatgatttacactggtcgtatgttgtaacataataagaagt 2235
Db 1706 GCTTGTATGTGCACACATGATGATTACACTGGCTGATTTGGTCACTACATACATACAGT 1765
QY 2236 acaattacccaatctgggagagacaacatgagagatgagaaatcacacacttaagctgaat 2295
Db 1766 ACACTGTGTCAATAGTGTAAGAGCTTCAGAGATGGGAAATCATATCTTACTGGAAT 1825
QY 2296 gtggagaggaagagaaatcgcagaatctgtctcaaaaagatgtgagagagagcagatgc 2355
Db 1826 GTGGGAGGAGGAAGAAATTCACACTCTGAGTCGGAAGATTAAAGAGAGCAAAATGC 1885
QY 2356 gcaattctctgttctgtctcgtctcgaagagctccaaatgtctcaacatggtatg 2415
Db 1886 GCAATTTCTTGTGTGTGTATGGTTCTCAGGGAGTTCCAAATGTTCTACATGGCGATG 1945
QY 2416 aatactgcacacaaaaggggcaacaacatactatgcatgattcttaatcaat 2475
Db 1946 AATATGTGCACACAAGAGGAGGAACAACATACGTACTGCCATGACCATTTATGCAANT 2005
QY 2476 atttcgtgtgataaaa--aagaaacactctcgtatgtgcacagatctgtcctca 2532
Db 2006 ATTTCGCTTGGGATTAAGAGAGAACATCTCTGATTTTACAGATTTCTCCGCTCA 2065
QY 2533 tgaccaaatctcgaagagagcgagggcttgccttgagagacttccaagggccaac 2592
Db 2066 TGACCGAATTCGCAAAAGATGTGAATCTCTTGGCTTGAGGACTTCCGACTTCAAGAC 2125
QY 2593 ggcgcagatgcatgcatcagcctgagaaagcctgatatgtctgagaaataagcagatcg 2652
Db 2126 GGTGAATATGCGACAGGTATCAGCGCCGGGAAGCCTGACTGGTCAGAGGCAACCGATTCG 2185
QY 2653 ttgccttccatgaagaatgaagaacagggcgagatctatgtgtgcttcaaccacagcc 2712
Db 2186 TTGCTCTCACATGAAGACCAACCAAGGGAGATCTACTGGCTTCCTTCAACACAGTC 2245
QY 2713 actaacggcgctgtgtatgagctccagagcgcgagggcgcggtgtgagaaacgggtgtg 2772
Db 2246 ACCTTCGGGTGTGTGTGTGGCTTCAGAGCGCTCTGGGTTCGATGGAGGCGGGGTGTG 2305
QY 2773 acacagcaagacacacacatacgaactctccacagcagacatctacatgcagctc 2832
Db 2306 ACACGGCAAGAGGACACATATGACTCTCCACCGATGGCCTGCACAGATCCTGCTGTCA 2365
QY 2833 ccatacacagctctcgaatctctctactccaaaccttaaccatlgctcagctcaat 2892
Db 2366 CCGTCTACCACTCTCTCATTTCCCAACTCCCAATCTCTATCTGCTACAGCTACTCCT 2425
QY 2893 cggcatctctatgtattgtgcacctgatgtttgagagagcaataata 2938
Db 2426 CCATCATCTCTGTATGTGGCCCTGATGTGAAGAACAGATATCA 2471
```

```
RESULT 2
US-09-187-124-1
; Sequence 1, Application US/09187124A
; Patent No. 6255563
```

```
; GENERAL INFORMATION:
; APPLICANT: Emmemmann, Michael
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES CODING FOR DEBRANCHING ENZYMES
; FILE REFERENCE: GFB8
; CURRENT APPLICATION NUMBER: US/09/187.124A
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: PCT/EP97/02292
; EARLIER FILING DATE: 1997-05-06
; EARLIER APPLICATION NUMBER: DE 196 18 125.9
; EARLIER FILING DATE: 1996-05-06
```

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1819)
; OTHER INFORMATION: Clone: IS05
US-09-187-124-1
```

```
Query Match 19.2%; Score 575.4; DB 4; Length 2133;
Best Local Similarity 69.8%; Pred. No. 9,5e-126;
Matches 799; Conservative 0; Mismatches 331; Indels 15; Gaps 1;
```

```
QY 600 gtaatggttccagcgctgtgtaacatctgcctcagctcagatgagctgcgatccctc 659
Db 1 gaattcgcacgagggccagagatgattgttgcctcccaatgycagatggtatcttc 60
QY 660 tccatatagcagtttgatgttggaagggcagcttaccttaagatatccctcaaaagacct 719
Db 61 tgcctctgacagtttgatgttggaagagagatctatctcgaagtttccacagagagatc 120
QY 720 ggtaatatagagatgcactgtgtggtatccagaagcatgtaaccagaatgtagaaca 779
Db 121 tgaatctatgaaatgcatgctgtgtgatttcaaaatcatgtagtgcgtgtaaacaaata 180
QY 780 tccgggttacttctatgtgagctgtgtcgaagcttgactatctgaagagctgagttaa 839
Db 181 tccgtgtacttccactgtgtgtgtgtggaagaaacttgatccactgaagaaacttgg 240
QY 840 tctgattgaatgaatgcccctgcagatgagttcaacagctggaagtaactcaacctcttc 899
Db 241 cgtatagagcttaatgcccctgcagatgagttcaacagctggaagtaactcaacctcttc 300
QY 900 -----caagatgaactcttgggatatcttaccataaactcttctacc 944
Db 301 tgtattggcgagctacaagtttaactcttgggatatcttaccataaactcttctacc 360
QY 945 aatgacaaagatcacatcagcgcgagataaaaactgtggcggtgtagccataaagatt 1004
Db 361 aatggaaagatcctgtctgcgtgtgtgaatgttcgagcttcgagccttcgacataaagatt 420
QY 1005 caaaacttttgaagaagagctccaaaagcggaatttgaggttcgttcgtgtctt 1064
Db 421 taagatctctgccaaggaagacacataaacgtggaacgaggtatcaatgagatgtctt 480
QY 1065 caacatacagcttgagggtaagatgagttcacaatatcatcttaagggggtcgatata 1124
Db 481 caatacactgttgaaaggaatgaaatgttccatacatacatcttaaggagcatgacaa 540
QY 1125 tactacatactatgatgttgacccaagggagagtttataactatctgtcgtgtggaa 1184
Db 541 cagttgttttatacagctgacctcaaggtgtaatttacaactactccagatgtgaaa 600
QY 1185 tactcttaactgtataatccctgtgtgtcgttaacttcatgttagattgtttaagacg 1244
Db 601 tactctcaactgtataatcccatgttagctgaatcttaagtgatgtgttgaagatgtg 660
QY 1245 ggtgacggaagcatgtgtatgtgttgccttgatccttgatccataatgacagagag 1304
Db 661 ggtacccgaagacagatgagatgtgtcgtcgtgtgtgtgtcttcatcttcaagaag 720
QY 1305 ttccagctgtgtggaatcgaatgagtgatgagagcccaatagaggtgacatgacac 1364
Db 721 tagcagctcgtgtgaaatgctgataatgtcatgtgaattcgaatgagcggtgacatgac 780
QY 1365 aacagaggaagcctctgttaccacacacactttatgacatgacagaaatgacacacatt 1424
Db 781 cacagagcctcctcacaagcccaactgattgataatgacatgacacacacatt 840
```





```

: PRIOR APPLICATION NUMBER: US 60/130,560
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: US 60/122,620
: PRIOR FILING DATE: 1999-03-03
: PRIOR APPLICATION NUMBER: US 60/119,386
: PRIOR FILING DATE: 1999-02-10
: PRIOR APPLICATION NUMBER: US 60/109,401
: PRIOR FILING DATE: 1998-11-20
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 71989
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic construct
US-09-443-501A-2

Query Match      2.08; Score 59; DB 4; Length 71989;
Best Local Similarity 49.6%; Pred. No. 0.00064;
Matches 181; Conservative 0; Mismatches 180; Indels 4; Gaps 1;

QY 2 gtcggggcggcgccgctgagcgggtggcgagcccaatgcagcgt---gcggggaag 57
DB 985 gtcgggctcttcttcagagcgctctgagtggtgaggaatgcgctgacgacgagcg 926

QY 58 gggctcggcggagtgctgagcgggtgtgagggcgagcgaagtagaggaagggg 117
DB 925 gctgctgctggcgccggcgagcgagtgatgagactgagagcgcgagtgctgagcg 866

QY 118 aggaagagagcggcgtgagcggaggaagtagcgcgtcggcgcgctgaggggtgctcg 177
DB 865 cggcgctgctacacttcgagcggcagccagctgagtgatggtgctgctcagacactcg 806

QY 178 cgggaatggccggcgccgtgagggcgccagccggcgctcgccgggggttcaatttcgcgcgt 237
DB 805 ACCAGGCGCTTCTGACAGCTCGCGCGCCGACGCGGCTGAGACAGTGCAGAGCCGGCG 746

QY 238 attcggcggagcagccagcggcgctctgctcttcacagcagaaatcgaagggcg 297
DB 745 ATGCGCGTGGCGAGCTATGCGCGCGCTCGCGGAGATGCGCAGAGAAATGCTGCGTCTG 686

QY 298 tgggggtgctcccgagtagagttcaccgcttgcgtgagcgcgccctttttgg 357
DB 685 CGGGCGCGCGCGCGGAGTCAAGCGCGCGGTGCTCGCGCGCGCGCGCTTCTGCTCG 626

QY 358 ggcctg 362
DB 625 ACCAG 621

RESULT 8
US-08-118-200-1
: Sequence 1, Application US/08118200
: Patent No. 6197500
: GENERAL INFORMATION:
: APPLICANT: SUTHERLAND, Grant R
: APPLICANT: RICHARDS, Robert I
: APPLICANT: SCHLESINGER, David
: APPLICANT: NAGARAJA, Ramaiah
: APPLICANT: KREMER, Eric J
: APPLICANT: YU, Sui
: APPLICANT: BAKER, Elizabeth
: APPLICANT: MULLEY, John C
: APPLICANT: MANDEL, Jean-Louis
: APPLICANT: PRITCHARD, Melanie April
: APPLICANT: LINCX, Michael
: TITLE OF INVENTION: DNA SEQUENCES RELATED TO ISOLATED
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
```

```

: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/118,200
: FILING DATE: 09-SEP-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/802,650
: FILING DATE: 05-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/672,232
: FILING DATE: 20-MAR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/638,518
: FILING DATE: 04-JAN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/966,517
: FILING DATE: 23-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Crane-Feury, Sharon E
: REGISTRATION NUMBER: 36,113
: REFERENCE/DOCKET NUMBER: 020160-164
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1028 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-118-200-1

Query Match      1.88; Score 53.8; DB 4; Length 1028;
Best Local Similarity 57.4%; Pred. No. 0.0017;
Matches 97; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 3 tcggggcggcgccgctgagcgggtggcgagcccaatgagcggcgaggagggtc 62
DB 327 TCTGAGCGGGCGCGCGCGGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCG 386

QY 63 ggcgagtgctgagcggcgtgtcagggcgagcgaagtagaagcagggggagg 122
DB 387 GCGGGCGGGCGCGCGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446

QY 123 gacgaagccggtggcgaggaagtagcgcgtcgggcgcgtagcagg 171
DB 447 GCGGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 495

RESULT 9
US-08-458-745-1
: Sequence 1, Application US/08458745
: Patent No. 6242576
: GENERAL INFORMATION:
: APPLICANT: SUTHERLAND, Grant R
: APPLICANT: RICHARDS, Robert I
: APPLICANT: SCHLESINGER, David
: APPLICANT: NAGARAJA, Ramaiah
: APPLICANT: KREMER, Eric J
: APPLICANT: YU, Sui
: APPLICANT: BAKER, Elizabeth
: APPLICANT: MULLEY, John C
```





```
ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
:
: NUMBER OF SEQUENCES: 7
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/1.797
:   FILING DATE: 08-JAN-1993
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: 749,621
:     FILING DATE: 19-AUG-1991
:     APPLICATION NUMBER: 224,114
:     FILING DATE: 25-JUL-1988
:
: SEQ ID NO: 1:
:   LENGTH: 3336
: 5457037-1
:
Query Match      1.8%; Score 53.2; DB 6; Length 3336;
Best Local Similarity 49.4%; Pred. No. 0.0039;
Matches 315; Conservative 0; Mismatches 278; Indels 45; Gaps 5;

QY 707 ctcaaaagagccggtgaataatagagatgacttgcgtgattaccgaagatga---tc 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1542 cgcagaagatgagtgatctacagagtgcatgycgcgcttaccgcagagacact 1601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 764 caagcaatgtagaacatccgggtacttcatctgagctgctgcgaagcttgactatga 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1602 ctatccctgcgcagatcgcgacactatacgggtgcagggcctcaagcagttacctg 1661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 824 aggaacttgagtgatctgattgaattaaagccctgcacgtgagttcaagcagctgagt 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1662 ccagcctgcgcgagcgcggtggaattcctgcgcgtgcaggaacgcagaatgctgca 1721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 884 ac-----tcaacctctctcccaagatgaactttgggagatctaccataact 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1722 acgattggtgtcccaattcgaagaccacaggaacttgggtacatgaccgagaact 1781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 935 tctttcaccaatgacagaatcacatcagcgagataaaaactgtggcgctgattgcca 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1782 acctctgcgcgagtcgcgctatgct-----acaaagcgcgctggcgctccca 1832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 taaatgagttcaaaactttctgaagagagctcacaacacgggagatgagttgactcgg 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1833 cgcgcgaggtccagcgatggtgtagcggtcttccacacgcagagatcaaggtctacatg 1892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1055 atgtgtcttcaaccatacagctgaggg-----taatgagaatg 1093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1893 atgtgtcttcaaccatacagctgaggg-----taatgagaatg 1952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1094 gtccaatatcatatttaaggggtcgaataactatactatactgcttgacc-----ca 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1953 ccccaattctatctggtgcgcggttggaacataccacgtactacagactcgacctcgggca 2012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1151 agggagagtttaataacttctgctgtggaataccttcaactgttaactatcctgtg 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2013 accaatattcttaagacaacacggcgtggtgcggaacttcaactgaactacacacggtg 2072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1211 ttcgtcaatcatctgtagatgttttaagatactggtgagcgaatgcatgtttatggtc 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2073 cgcagaacctatcgtctgactgctgtgcatctggtgcgaacacgattggtcggtgattgct 2132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1271 ttcgtttgactctgcatcataatgacagaggtcc 1308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2133 ttcgcttgagactgtcttcggtgctgcgacaagttgc 2170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-576-626A-2
: Sequence 2, Application US/08576626A
: Patent No. 5998194
:
GENERAL INFORMATION:
: APPLICANT: Summers, R.G.
: APPLICANT: Katz, L.
: APPLICANT: Donadio, S.
: APPLICANT: Staver, M.J.
```

```
TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
:
: TITLE OF INVENTION: BIOSYNTHESIS GENES
:
: NUMBER OF SEQUENCES: 60
:
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Abbott Laboratories
:   STREET: 100 Abbott Park Road
:   CITY: Abbott Park
:   STATE: Illinois
:   COUNTRY: USA
:   ZIP: 60064-3500
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTED Version 2.0
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/576,626A
: FILING DATE: 21-DEC-1995
: CLASSIFICATION: 435
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
:   NAME: Diane Casuto
:   REGISTRATION NUMBER: P-40,943
:   REFERENCE/DOCKET NUMBER: 5857.US.01
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: (847) 938-3137
:     TELEFAX: (847) 938-2623
:   TELEX:
:
INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 8051 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:
US-08-576-626A-2
:
Query Match      1.8%; Score 52.8; DB 2; Length 8051;
Best Local Similarity 53.6%; Pred. No. 0.0071;
Matches 133; Conservative 0; Mismatches 112; Indels 3; Gaps 1;

QY 74 cgcgcggtgtgtcgaagcgcgcagagtagaagacgaagggagaggaagcagcgct 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 CGCGGGGCTGGCGCCGACACGCTGTCACGTGACAGAGTGCCACCGCGTTCGCCGCCGC 753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 134 ggcggagagacaggtacgcgt---cggcggcggtgcaggtgtctgcgggaatgcgcgc 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 GCTGAGACACACACACGCGCTGGCGCGGCGGACGTGGCGCTGGCGCGGACCGATCCGA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 gccgctgggcgcacacgcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 GCCGCTGGGCGACATCTTCGCGCGCGTCTCCGCGAGCGTGCAGCGGACGAGCGGAGCC 873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 251 caccgcggcggtcgtctgcctcttcaagccagaagatctcaagcgtgtggtgtcctcc 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 CGCGGTGACGTGTGTCACCGTGCCTCCGCGCGGACGACGCGGACGACGACGATTCGCGAG 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 311 cgaataga 318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 934 CGACGACA 941
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-721-822A-10
: Sequence 10, Application US/09721822A
: Patent No. 6306606
:
GENERAL INFORMATION:
: APPLICANT: Michael J. Weber
: APPLICANT: Jacqueline Wyatt
: APPLICANT: Lex M. Cowsett
: TITLE OF INVENTION: ANTISENSE MODULATION OF MP-1 EXPRESSION
```

FILE REFERENCE: RTS-0142  
CURRENT APPLICATION NUMBER: US/09/721,822A  
CURRENT FILING DATE: 2000-11-22  
NUMBER OF SEQ ID NOS: 135  
SEQ ID NO 10  
LENGTH: 1315  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (147)...(521)  
OTHER INFORMATION:  
US-09-721-822A-10

Query Match 1.7%; Score 52.4; DB 4; Length 1315;  
Best Local Similarity 79.5%; Pred. No. 0.0041;  
Matches 62; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2920 ttgagagaccatatacagtaataatgctctatatgtcaaaaaaaaaaaaaa 2979  
DB 1221 ttcttaccatgctgtaataaataagtagtatttaaggtcaaaaaaaaaa 1280  
QY 2980 aaaaaaaaaaaaaa 2997  
DB 1281 aaaaaaaaaaaaaa 1298

RESULT 14  
5457037-2  
PATENT NO. 5457037  
APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;  
GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO  
TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE  
ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME  
NUMBER OF SEQUENCES: 7  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/1,797  
FILING DATE: 08-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 749,621  
FILING DATE: 19-AUG-1991  
APPLICATION NUMBER: 224,114  
FILING DATE: 25-JUL-1988  
SEQ ID NO: 2  
LENGTH: 2253  
5457037-2

Query Match 1.7%; Score 51.6; DB 6; Length 2253;  
Best Local Similarity 49.2%; Pred. No. 0.0079;  
Matches 314; Conservative 0; Mismatches 279; Indels 45; Gaps 5;

QY 707 ctcaaaagagacctgtatataatgagagcacttcgtgtgattcagaagcatga---ct 763  
DB 524 cgcgaagaagatgtgtactcagaggtgctgcgcgcctccagcagagacact 583  
QY 764 caagcaagtagaacatccgcgtactcttcattcagagctgtgcagagcttgacttga 823  
DB 584 ctatcccgccgagatccgcgcgcacttaacggtgcagaggtctcaagccaattactcg 643  
QY 824 agagagctgagtaattgtatgaatgaatccctccagcatagttcaagcagctggaat 883  
DB 644 ccaagcctggcgctgacgcggtggaattccctgcggtgcaggaagcagaatgatgcga 703  
QY 884 ac-----tcaaccctctctccaagatgaacttttggtgatatctaccataact 934  
DB 704 acgagtgtgtctccatccagatgcccacacgaactactcgtggctacatgacgagaact 763  
QY 935 tctttcccaatgacacatcagcagcggtataaaaaactgtggcgctgtagtgcga 994  
DB 764 actctgcgcgagatccgcgcctatgcct-----acaacaagcgcgctggtggtccca 814

QY 995 taaatgattcaaaaactttgttaagagaggtcacaaaacgggggaattgagtgatcctgg 1054  
DB 815 cggcgagttccagcgatgtgtcaggcglttcacaaacgagcatcaaggctactacatgg 874  
QY 1055 atgtgtctcaaacacacagcgcgaggg-----taatgagatg 1093  
DB 875 atatgtctcaaacacacagcgcgaggg-----taatgagatg 934  
QY 1094 gtccaatatcatcatttaaggggtgcataatactatactatgctgcacc---ca 1150  
DB 935 ccacattattatcgtggcggtgtgacaataccacgtactacgagctgactcgggca 994  
QY 1151 agggagagtttaactattctgcgtctgtgggaatcactcaactgaatcatcctcgtgg 1210  
DB 995 accaatattctcagcaacacagcgatctggcggaactcaatacgtacacacaggtgg 1054  
QY 1211 ttgcataatcatctgtagattgtttaagatcaggtgaggaatgcatggttagtgtt 1270  
DB 1055 cgcagaaccttaacgtcgcagctgcgcgtatctggcggaacagatggcggtgagtgct 1114  
QY 1271 ttgcattgactctgcatcacaataatgacagaggtcc 1308  
DB 1115 ttcgctcgacctgcttcgctgcgtcgcgaacaggttgc 1152

RESULT 15  
US-09-056-556-171  
Sequence 171, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 535 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-171

Query Match 1.7%; Score 50.8; DB 4; Length 535;  
Best Local Similarity 52.2%; Pred. No. 0.0065;  
Matches 142; Conservative 0; Mismatches 122; Indels 8; Gaps 1;

Db 2 CCGGCGCCACCGGGGACCGGGTTCCGCCGTGCGCGCGGGGCGCGCGGACAGGGCG 61  
QY 69 gtcgtcgccggtgtgtcgagcgagcgaaggtagaagcgaagggagagagcgaag 128  
Db 62 GTATCAGCGCGTCCCGCGGACCAACGGCTCTGCGTGGCGCGGACCGCGGACAG 121  
QY 129 ccggtgagcgaagcaggtacggtcg-----cggcggtgcaagggtgtcgccg 180  
Db 122 GCGGCGCGGGGCGCTGCGCGGGCGCGCGCGATTAACCCACCGGCAATCGCGCGCG 181  
QY 181 gaatgccgagcggtgtgtcgccacgagcggtcgagcggtcatttcgcgtctatt 240  
Db 182 GCGGCAACCGCGCGGACCGGGAGCGGCCGCGCGCGCGGTGCGCGCATCGGTA 241  
QY 241 ccggtgagcgaagcagcggtcggtcggtcgtcgc 272  
Db 242 CCGGCGGACCGCGCGCGGTGCGGACAGCGTC 273

Search completed: July 31, 2002, 13:03:22  
Job time: 17056 sec





<b>RESULT 3</b>					
LOCUS	BG444934		913 bp	mRNA	linear EST 15-MAR-2001
DEFINITION	GA_Ea0056B23f Gossypium arboreum 7-10 dpa fiber library Gossypium				
ACCESSION	BG444934				
VERSION	BG444934.1		GI:13354586		
KEYWORDS	EST.				
SOURCE	Gossypium arboreum.				
ORGANISM	Gossypium arboreum.				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.				
AUTHORS	1 (bases 1 to 913) Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A. An integrated analysis of the genetics, development, and evolution of the cotton fiber Unpublished (2000)				
TITLE					
JOURNAL	Contact: Wing RA				
COMMENT	Clemson University Genomics Institute				
	Clemson University				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Tel.: 864 656 7288				
	Fax: 864 656 4293				
	Email: rwing@clemson.edu				
	Seq primer: TAATAGCATCCTATAGGG				
	High quality sequence stop: 777.				
FEATURES	location/Qualifiers				
Source	1..913				
	/organism="Gossypium arboreum"				
	/strain="AKA"				
	/cultivar="8400"				
	/db_xref="taxon:29729"				
	/clone="GA_Ea0026B23f"				
	/cclone_lib="Gossypium arboreum 7-10 dpa fiber library"				
	/issue_type="Fibers isolated from bolls harvested 7-10 dpa"				
	/lab_host="E. coli"				
	/note=Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"				
BASE COUNT	254 a	158 c	235 g	266 t	
ORIGIN					
Query Match	13.0%;	Score 390.2;	DB 10;	Length 913;	
Best Local Similarity	72.2%;	Pred. No. 1.9e-34;			
Matches 542:	Conservative	0;	Mismatches 193;	Indels 16;	Gaps 2;
OY	586	ataacgcaaggagggatgatgcgtcttcacagcgctggttaacaattgctggccctcagaatgct	645		
Db	26	ATAAGCAGAGGGGAGATTGGGGCTTTAGGGCGCTTGANAGTAATTTGTGGCCCCAAATGGCC	85		
OY	646	ggcatgatccctcttccatatagcacgcttgattggaaggcgacctacccttaagalat	705		
Db	86	GGAATGGGCCCTACTTCAGAATATACAGTTGATTTGGGAAGGATGTACCTTGAGACAT	145		
OY	706	cctcaaaaggagcccggttatatargagttagactgctggtgatctacagaagcatgatcca	765		
Db	146	CCACACAGAGATCTCATATATTATTAGCAAAATGCATGCGCGATATATACAAGCATGAATCT	205		
OY	766	agcaaatgtagaacatccgcggtactcttcatctggagctgtctcgaagcttgactattgaag	825		
Db	206	AGTGGGACAAATTTCTCCGCGCACGATNCGTGTGTGTGTGGTGGAAAAGCTTGACCATTAAAG	265		
OY	826	gaagcttgagagtaaattgctatcgaattaaigccctgcctcgaatgaatccaagcagctggaatc	885		
Db	266	GAACCTTGAGATCAACTGCATAGAAATTAATGACCATGTCAAAGATTCATGAGATGAGATAC	325		
OY	886	tcaaccctcttc-----caagatgaaccttttggggatattctaccta	930		
Db	326	TACAGCTACAAATTCGTTTTGGGTGACTATAGATTTGAAGCTTTTGGGATATTCACCAATTC	385		

QY 931 aactctcttcacaaatgcagaaatgcacatcaggcgagatataaaaaactctgctggcgat 990  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 386 AATTATTTTTCGCCCATGATATAGGTATTTCATCTTGSTATTCTGACTGTGGTCGTAT 445  
 QY 991 gccataaatgatgtccaaactcttgtaagagaggtcccaaacggggaaattgagtgatc 1050  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 446 GCATTTAATGAAATTCATAATATCTTTAAAGAAAGCACATTAACCGGAATTGAGGTTATA 505  
 QY 1051 ctggaatgtgtctcccaaccatacagctgtaggtaatgaaatggtgccaatatcatcatt 1110  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 506 ATGGATGTGGTTTTCATTCACACAGCTGAAGCGCATGGAAGGCTCAGTTGTGCATTTT 565  
 QY 1111 aagggtgtcgataatactacataataatgtctgacccaagggaagattataaacat 1170  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 566 AGAGGTGTGATTAACGTGTCTATTATTCATGCTGGCACCTTAAGGAGAGTACTACAAATTAT 625  
 QY 1171 tctgctgtgtggaataacctcaactgtcaatcatcctctgctgtcgtccaaatcatctgagat 1230  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 626 TCAGAGGTGTGGAAACACATTCATCACTGTGACCATCTGTGGTGGCTCAATTTATATTAGAC 685  
 QY 1231 tgtttaagaatactcyyggtgaaagaaatgcataatgttgaatgttctgttttgaatcttgacac 1290  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 686 TGCCCTAAGATTTTGGGTAAACAGAAATGCATGCTGATGGGTGCTTGTGATCTGTGCTTCC 745  
 QY 1291 ataatgaccagaggttccagctctgtggatc 1321  
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 746 AT-ATGACCAAGAGTAAAGCGTCTTTGGGATC 775

RESULT	4
LOCUS	BE422551/c
DEFINITION	BE422551 435 bp mRNA linear EST 24-JUL-2000
ACCESSION	WHE00055.F11_K212S wheat endosperm cDNA library Triticum aestivum
VERSION	BE422551
KEYWORDS	BE422551.1 GI:9420310
SOURCE	EST.
ORGANISM	wheat.
REFERENCE	Triticum aestivum
AUTHORS	Enkayotz; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
TITLE	1 (bases 1 to 435).
JOURNAL	Altendach, S., Anderson, O. D., Cho, S., Gall, G., Han, P. S., Hsia, C. C., Kang, Y., Lazo, G. R., Miller, R., Rausch, C. J., Seaton, C. L. and Tong, J. C.
COMMENT	The structure and function of the expressed portion of the wheat genomes - Endosperm cDNA library
	unpublished (2000)
	Contact: Olin Anderson
	US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
	800 Buchanan Street, Albany, CA 94710, USA
	Tel: 5105595773
	Fax: 5105595818
	Email: oanderson@pv.usda.gov
	Sequence has been trimmed to remove vector sequence and low quality sequence with phred score less than 20
	Seq primer: Stratagene SK primer.
FEATURES	Location/Qualifiers
SOURCE	1..435
	/organism="Triticum aestivum"
	/cultivar="Cheyenne"
	/db_xref="taxon:4565"
	/clone="WHE0055.F11_K21"
	/clone_1lb="wheat endosperm cDNA library"
	/tissue_type="Endosperm"
	/dev_stage="5 to 30 days post anthesis seed"
	/lab_host="E. coli SOLR"
	/note="Vector: Lambda ZAP II, excised phagemid; Site_1: EcoRI. Seeds collected, endosperm isolated, and RNA prepared by Susan Altendach. Library constructed by Stratagene, Inc. Plasmid DNA preparations and DNA







Db 556 TAATATCCATTGTACGTCAATTATATGATGGATTG 591  
 RESULT 8  
 B1934559 580 bp mRNA linear EST 18-OCT-2001  
 LOCUS EST554448 tomato flower, anthesis Lycopersicon esculentum cDNA  
 DEFINITION clone cTOD2005 5' end, mRNA sequence.  
 ACCESSION B1934559  
 VERSION B1934559.1 GI:16249031  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Lycopersicon esculentum; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 580)  
 van der Hoeven, R.S., Bezzerides, J.L., Karanymcheva, S.A., Tsai, J.,  
 Utecherback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M.,  
 Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
 Generation of ESTs from tomato flower tissue, anthesis (2001)  
 Unpublished (2001)  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 EMAIL: <http://www.genome.clemson.edu/orders/index.html>  
 THIS CLONE IS AVAILABLE THROUGH THE CLEMSON UNIVERSITY GENOMICS  
 INSTITUTE  
 SEQ PRIMER: T3.  
 FEATURES  
 source location/Qualifiers  
 1..580  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA96"  
 /db\_xref="taxon:4081"  
 /clone="cTOD2005"  
 /clone\_lib="tomato flower, anthesis"  
 /tissue\_type="flower"  
 /dev\_stage="anthesis"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Cornell University; sequencing: The  
 Institute for Genomic Research; flower buds and flowers  
 were taken from greenhouse plants (4-8 wks old, TA96).  
 They were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."  
 BASE COUNT 168 a 100 c 125 g 187 t  
 ORIGIN  
 Query Match 10.1%; Score 303.2; DB 10; Length 580;  
 Best Local Similarity 72.9%; Pred. No. 8.7e-25;  
 Matches 411; Conservative 0; Mismatches 138; Indels 15; Gaps 1:  
 QY 672 gttgattgggaaggcgaactcctaagatactcaaaaggaagctgtaataataga 731  
 |||||  
 DB 16 ctttgattgggaaggagatctacacactgaagtttctcagagaaatctgtaattgatga 75  
 QY 732 gatgacttgctggatcaacgaacgatgatcaagcaaatgtagaacatccgggtactt 791  
 |||||  
 DB 76 AATGCATGTGCTGATTTACAAATCATGAGTCGAGTGAACAAATATATCTGTGTAATTA 135  
 QY 792 catggaagctgtgctgaactatattgaaggagcttggaatgaatgtaattgaat 851  
 |||||  
 DB 136 ccttgctgttggaagaaacttgatcacttggaagaaacttgctgtaactgtaagact 195  
 QY 852 aatgcctgcacagatcaacgaacgtggaactcaactctcttc----- 899  
 |||||  
 DB 196 AATGCCCTGTACGAGTCAATGAGCTGAGTACTTATTAATCTGTATATGGGCA 255  
 QY 900 ---caagatgaactcttgggagatattcaacataaactcttccacaaatgaagata 956  
 |||||

Db 256 CTACAGTTTAACCTTTGGGGCTATTCTACTGTCAATTTCTTTCTCCAAATGGAGAATA 315  
 QY 957 cacatcaggcggagataaaaactgtggcgatgacatcaatagatgacaaacttgt 1016  
 |||||  
 DB 316 CTCACTCTGCTGCTCACTAATTTGAGCTCGGTCAATAACGAATTAATGATCTTGT 375  
 QY 1017 aaggaaggtccacaaagggaagatgagtgatccctgagatctgtctcaaccatacagc 1076  
 |||||  
 DB 376 CAGGAGACCAATTAACCTGGAAATCAGATGATATCATGTGATGTTTTCATACACATGCG 435  
 QY 1077 tgaagtgatgagaatgtgtccataatcatcttaaggagggtcgaataactataacta 1136  
 |||||  
 DB 436 TGAAGGAATGAATATGATCCCATCTATCAATTTAGAGCAATGACAAAGTGTTTTA 495  
 QY 1137 tatgttcaccccaaggagagatttaactatctcgtgtgggaatccctcaactg 1196  
 |||||  
 DB 496 TACGCTACCTCCTTAAGGCTGAATTTTACAACTACAGATGTGAAATACCTTCACTG 555  
 QY 1197 taatcatcctgtgtgtcgaatt 1220  
 |||||  
 DB 556 TAATATCCATTGTACGTCAATT 579  
 RESULT 9  
 AM832588 558 bp mRNA linear EST 03-DEC-2001  
 LOCUS sm14b07.y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1027-8846 5' similar to TR:022637 022637 S01 ISOMYBLASE. ;  
 mRNA sequence.  
 ACCESSION AM832588  
 VERSION AM832588.1 GI:7926562  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 558)  
 Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Coryell, V., Rhanna,  
 A., Bolla, B., Merritt, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
 T., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schick,  
 R., Ritten, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann,  
 R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 CONTACT: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@wustl.edu](mailto:est@wustl.edu)  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: [ccu@resgen.com](mailto:ccu@resgen.com)  
 High quality sequence stop: 411.  
 FEATURES  
 source location/Qualifiers  
 1..558  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1027-8846"  
 /clone\_lib="Gm-c1027"  
 /tissue\_type="cotyledons of 3- and 7-day-old Williams  
 seedlings"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; this cDNA library was constructed from mRNA isolated  
 from cotyledons of 3- and 7-day-old Williams seedlings  
 which were propagated on paper towels with distilled  
 water. The cotyledons were flash-frozen in liquid  
 nitrogen, then lyophilized for 72 hours. Unequal amounts

of mRNA was used for cDNA synthesis. Stratagene's cDNA Synthesis Kit (Catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (V-A, C, or G) was added to the 3' end of the primer (GAGACAGACAGACAGACAGACTGCTCCAG(T)18) to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript(II) M13XP1 digested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell.

BASE COUNT 156 a 127 c 111 g 164 t  
ORIGIN

Query Match 9.3%; Score 278.6; DB 9; Length 558;  
Best Local Similarity 70.0%; Pred. No. 4.6e-22;

Matches 390; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

OY 2349 cagatgacgaattcttctgtctcgaatgttccaaagagttccaaatgtctacatg 2408  
DB 1 CAAATGCGGAATTTTCTTCTCTCATGTTTCCAGGAGATTCGAATGATATATG 60  
OY 2409 gttgatgaatattggcacaacaaaggaggacacaacatactatgcatatctat 2468  
DB 61 GCGGATGAATATGGACACACAAAGAGGAAATACATCTTGTCTGATATATAT 120  
OY 2469 gtcattatcttcgtcgtatataaaagacaataa--ctgtagttgcacgattctg 2525  
DB 121 CATTAATTTACTTCCAAATGGGACAAAGAAAGAAATCCATCAAGCTTTCAGATTTTGT 180  
OY 2526 tgcctcatcaccaaatctcgaagagtgagaggtctgtgaccttgagacttccaa 2585  
DB 181 GCGCTTATGACTAAGTTCGCCGACGAATGGAATCCGTAGCTTACGCTGCCAAC 240  
OY 2586 gccaaacggcgcagatgagatgcatcagccttggaagccctgattgtctgaataag 2645  
DB 241 TCTGAGAGCGCTGACAGTGGTCTCATTTCTCGAAGACGACAGCTGTAACACG 300  
OY 2646 cgattcgttgcctcttccttgaaagatgaaagcagggagatctatgtgccttaac 2705  
DB 301 CTTTGTGTGGCTTGTACCATGTAATTCAGTAAGGGAATAATATACATTTGCTTAAT 360  
OY 2706 accagaccacttaccggcgtgtgtgagctcccgagagcgcgcggttggaacg 2765  
DB 361 ATGAGTCATTACCTTTCACAGTTACCTTCCGAGACGCTCTGATACAAATGGAACT 420  
OY 2766 gtgtgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2825  
DB 421 CTTGTAGACACACGAAAGCTACCATATGATTTCTCTACCTCTGACCTTCTGGAAGA 480  
OY 2826 gcttcacacataacaggttcgcatcttcctactcaactcaacccatgctcagc 2885  
DB 481 GATATGTCATACACAGATGCTCAGTTCTGAGACGCAATATGATATCCATGCTTAGA 540  
OY 2886 tactcatcggtcatcct 2902  
DB 541 TATTTCTTCATATATCT 557

## RESULT 10

BM178903

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

TITLE

JOURNAL

COMMENT

BM178903 547 bp mRNA linear EST 06-DEC-2001  
sa360h07.y1 Gm-c1072 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1072-4213 5' similar to TR:022637 02637 SU1 ISOMYLAASE. ;  
mRNA sequence.

BM178903.1 GI:17402121

EST.

soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 547)

Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V., Khanna

A., Bolla, B., Maitra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyle, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pope, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., Mccann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R./Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact: c@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 422.

Location/Qualifiers

1. 547

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1072-4213"

/clone\_1ib="Gm-c1072"

/tissue\_type="seedlings induced for symptoms of SDS

(Sudden Death Syndrome) disease"

/dev\_stage="2-3 weeks old"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:

XhoI; The cDNA library was constructed from mRNA isolated

from 2-3 week old seedlings that were induced for symptoms

of SDS (Sudden Death Syndrome) disease by the

translocation of culture filtrate of Fusarium solani f.

sp. glycines (Plant Cell Report 18:375-380). Cultivar PI

56374 is partially resistant to the disease SDS. Plant

tissue (expanded leaves, folded leaves, and new shoots)

were collected at 1, 6, 24, and 48 hrs. after inoculation

and their mRNA pooled equally for cDNA construction. The

library was prepared using the Stratagene pBluescript II

SK(+) library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with an XhoI restriction site. EcoRI

adaptors were ligated to the blunt-ended cDNA fragments

BASE COUNT

151 a 124 c 111 g 161 t





Oy	2816	accatgagcgtccacacataacaggcttcgcgatctccactccaacctaacc	2875
Dd	316	NCCAGATGCTGCTGCACCGCTTACCATTCTCTCTCCTCAACTCATCTCTATCC	257
Oy	2876	cagctcagctactcatcgatcattagtagtatgtcgccctgatgtttagaga	2928
Dd	256	TATGCTAGCTACTTCCTCATCATCTGTATGTGGCCCTGATGCTGAAGA	204
<hr/>			
RESULT	14		
LOCUS	BEA75829/c	472 bp	mRNA linear EST 28-JUL-2000
DEFINITION	946048H01.x2 946 - tassal primordium prepared by Schmidt lab Zea		
ACCESSION	mays CDNA, mRNA sequence.		
VERSION	BEA75829		
KEYWORDS	BEA75829.1 GI:9566320		
SOURCE	EST.		
ORGANISM	Zea mays. Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 472)		
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946048 row: H column: 01. Location/Qualifiers		
FEATURES			
source	1..472 /organism="Zea mays" /cultivar="OH43" /db_xref="taxon:4577" /clone_lib="946 - tassal primordium prepared by Schmidt lab" /tissue_type="tassels" /dev_stage="just after the transition from vegetative to inflorescence development" /lab_host="XLDLR"		
BASE COUNT	125 a 106 c 128 g 113 t		
ORIGIN			
Query Match	7.6%; Score 227.6; DB 10; Length 472;		
Best Local Similarity	78.6%; Pred. No. 2, le-16;		
Matches	272; Conservative 0; Mismatches 74; Indels 0; Gaps 0;		
Oy	2593	ggctgcaagtgcacatgcacgcacctgggaagcctgatgtgctcgaataagcagatcg	2652
Dd	472	GGTTCGATGACGACGATCATGACGCCGGGAACCTTGACTGTCAGAGCCAGCATTTGG	413
Oy	2653	tggccttcctcgttaaaagttaaagaacaggggagagatgttgctcacaacacagcc	2712
Dd	412	TTCCTTCACCATGAAGAGGAGAAACCAAAGGCGAGATCTACGCGGCTTCACAACACAGTC	353
Oy	2713	acttacagcgcgttgttgatgcccccagacgcgcgaagggcgccggttggaaccggtgtgg	2772
Dd	352	ACCTGTGCGTGTGTCGTCGGGGCTTTTAGACGCGCTCTGGGTTCGANTGGAGCCGATGTGG	293
Oy	2773	acacaggaagccagcacataagacttccataccagcagacttactcgtatcgcgctcta	2832

Db	292	ACACGGAGAGAGGACACCATATAGCTTCCACATGATGAGCGCTGGCAGATCGTGGTGTCA	233
Qy	2833	ccatacaccagtlctgcgatttctctctactccaaacctctaccatgtcgaactcat	2892
Db	232	CCGTTACCAAGTTCCTCATTTCCCACTCACTCAATCTATCTATGTCACATATATCT	173
Qy	2893	cggtactcttagtatgctgcgcctgagtgttgagagaccatata	2938
Db	172	CCATATCTCTGTATGGCGCCCTGATGTGTGAAGAAAGGATATACA	127
RESULT	15		
LOCUS	BI120843	456 bp	mRNA linear EST 31-DEC-2001
DEFINITION	F023P78Y Populus flower cDNA library Populus balsamifera subsp. trichocarpa cDNA, mRNA sequence.		
ACCESSION	BI120843		
VERSION	BI120843.1	GI:18004818	
KEYWORDS	EST.		
SOURCE	Populus balsamifera subsp. trichocarpa.		
ORGANISM	Populus balsamifera subsp. trichocarpa		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.		
AUTHORS	1 (bases 1 to 456) Hertzberg, M., Aspeberg, H., Erlandsson, R., Bjorkbacka, H., Hiltonen, T., Karlsson, J., Teeri, T., Gustafsson, P., Bahlero, R., Jansson, S., Nilsson, O., Sundberg, B., Nilsson, P., Uhlen, M., Sandberg, G. and Lundberg, J.		
TITLE	Gene expression in Populus		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Erlandsson R Department of Biotechnology Royal Institute of Technology Teknikringen 30, Stockholm S-10044, Sweden Tel: 46 8 790 8287 Fax: 46 8 245452 Email: rikeri@biochem.kth.se.		
FEATURES	Location/Qualifiers		
source	1..456 /organism="Populus balsamifera subsp. trichocarpa" /db_xref="taxon:3694" /clone_lib="Populus flower cDNA library" /note="Organ: flower"		
BASE COUNT	150 a 85 c 119 t 3 others		
ORIGIN			
Query Match	6.8%; Score 202.4; DB 10; Length 456;		
Best Local Similarity	74.1%; Pred. No. 1.3e-13;		
Matches	335; Conservative 0; Mismatches 108; Indels 9; Gaps 6;		
Qy	2140	atcagcaggaaggaagaaaccttgcacagaatcaactttatgtgacatgtagt	2199
Db	6	ACCAGAAAGAGAGAAACCATGGAACAGCATCTTGTATGTGCATGATGATT	65
Qy	2200	ttacacatgctgatttgcgtacataataaagaagatataccaatg- ggaagac	2258
Db	66	TTACTTTGGCTGATTAGTGACATATAACAAAGACATATCTTGGCAATGGCGAAGCAC	125
Qy	2259	aacagagatgagaaa- atcacaatctagctcgtgaatgttg- ggaaggaaggaatctg	2316
Db	126	AACAATGATGAGAAACATCATACATATAGCTGGAATTGTGGCACAGAGTTGAATTGG	185
Qy	2317	caagatgtcgttcaaaaagaatggagaagaag- -cagatgcgaattcttgttctgtc	2374
Db	186	CCAGCATTTTCAGTGAAGAAAMNTTGCAGAAACAGACAAATGAGAAATTTTCTTCTGTGT	245
Qy	2375	catgtttctcaagaagatcacaatgtctacatggtgtgataatggtccacacaaag	2434
Db	246	CATGGTTTACAAAGGTGTCCCAATGATATACATGGGTGATGATGATGTCACACAAAG	305



```
OY 2435 gggcaacaacatcacatcgcctgacatgattcttcaatattctgcgtggataaaa 2494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 GGGAAACAA-AATACATATTGCCATGATACCCCTATATTACTACTTCGCGCTGATAGAA 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2495 agaacaatac--tctgagltgcacgcgattctgctgctcctcagcaaatlcgcaagga 2551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 GGAAGAATCCTCATACAGACTTCTTAGATTGGCCGCTCATGACCAAGTTCGCGCATGA 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2552 gtgcgagggtccttggccttgaaggacttccaa 2583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 ATGTGAGTCCCTTGTTGTTGAATGACTTCCCA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 31, 2002, 10:50:03  
Job time: 9127 sec





QY 224 VEHPGFIGAVSKLDYLKEELGAVNCIELMPCHEFNELEYSTSSKNNMNGXTINFPSPMT 283  
Db 277 tkhpqcyigaavskldhlyelgvncielmpchelefnelysssskmfwysctinfstis 336  
QY 284 RYTSGGIKNCGRDALNEKTEFVREAHKRGIEVLDVFNHNTAEENGPILSFKGVDMNT 343  
Db 337 ryssagldsgcgalnefkafvrehkrgylevmdvfnhtegnekprlslfrgdnst 396  
QY 344 YMLAPKGEFVNYSGCGTFCNCPNVROFIYDCRLRYWTEMHVDFRFDLASIMTRGSS 403  
Db 397 ymlapkgelynysgcgtfcncnbpvreflvdclrywtemhvdgfrfdlasiltgrcs 456  
QY 404 LMDVNVYGCALIEGDMITTPRLVPRPLIDMISNDPIILGAVKLIEMADAGLLYOGGPR 463  
Db 457 lmdpvnvyspmegmltctgprlvpplldmisdndpilaawdaggllyegqfpr 516  
QY 464 HMWVSENGKVRDVIROFIKTDGAFACLCGSPILYOGGRKPMHSINFCANHG 523  
Db 517 hmwvsewngkyrdvtrvrlfkygdgafagaefaelcgspilyaggrtkpmhsigfvcadng 576  
QY 524 FTLADLVYTKRKYMLPNCENNRGENHNLNMGCEGEFARLSVRLRKQRNRFVCLM 583  
Db 577 ftladlvynakylungedfrdgenhnlwncgegefaslsvrlltkrqmtnffvclm 636  
QY 584 VSGVPMFYMDEYHTGNGNNNTYCHDSYNYFRMDKK-EOYSELHRRCCMTFRKRC 642  
Db 637 vsqgvpmfymdeyhtgngnnntychdshyvyfrwdkkesgsdlyfrctimtefikcc 696  
QY 643 EGLGLEDPTAKRLQWGHOGPKPMSENSRFVAFSMKDERGELIVANFTSLDAVVEL 702  
Db 697 eslgedptcerlkwhnbgpkrpmsaasrfvafctmdekelyvalnctshlpvavgl 756  
QY 703 PERAGRWEPVVDGKPAFYDLTDLPDRALTIHQSHFLYSNLVPMLSYSSVILVLRP 762  
Db 757 peragfwrwepvvdtkesaydfldtgdldpravtyqfshflsnllypmlsyslilvltrp 816  
QY 763 DV 764  
Db 817 dv 818

RESULT 3  
AAY00869  
ID AAY00869 standard; Protein: 793 AA.  
XX  
AC AAY00869;  
XX  
DT 21-MAY-1999 (first entry)  
XX  
DE S. tuberosum isoamylase clone 15 protein sequence.  
XX  
KM Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;  
KM starch modification; debranching activity.  
XX  
OS Solanum tuberosum.  
XX  
PN MO9906575-A1.  
XX  
PD 11-FEB-1999.  
XX  
PR 30-JUL-1998; 98WO-G802280.  
XX  
PR 31-JUL-1997; 97GB-0016185.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Bustos Guillen R, Edwards EA, Martin CR, Smith AM;  
XX  
DR MPI: 1999-153803/13.  
XX  
DR N-PSDB; AAX27061.  
XX  
PT New nucleic acid from potato encoding starch debranching enzyme with  
PT isoamylase activity - and related vectors, transformed cells and

PT plants, proteins and antibodies, used to generate starch with  
PT altered properties, for use in foods, as thickeners etc.  
XX  
XX  
PS Claim 2; Fig 5; 78pp; English.  
CC This sequence is a Solanum tuberosum isoamylase of the invention.  
CC Fragments of the isoamylase coding sequence are used to identify and  
CC clone isoamylases from other plant species, by standard hybridisation  
CC or amplification methods. Expression of the isoamylase DNA in host cells  
CC is used to produce the isoamylase which is used: (1) to raise antibodies  
CC (Ab) for identification, isolation and localisation of isoamylases; and  
CC (11) for synthesis of branched polysaccharides. The DNA, proteins  
CC and Ab (or related peptides) are used to alter the quality and  
CC quantity of polysaccharides in a host cells, specifically to alter  
CC branching in amylopectin. The modified starches produced are useful in  
CC human or animal foods; as biodegradable plastic; as food or paint  
CC thickener; in starch-coated films, paper and textiles; in mining  
CC explosives; in pharmaceuticals and glues. Manipulation of debranching  
CC activity in a plant allows control of starch properties, e.g. increased  
CC gel strength; formation of paste rather than gel, changed physical  
CC characteristics etc.  
SQ Sequence 793 AA:  
  
Query Match 71.9%; Score 3029; DB 20; Length 793;  
Best Local Similarity 71.4%; Pred. No. 3.7e-260;  
Matches 546; Conservative 78; Mismatches 113; Indels 28; Gaps 7;  
  
QY 10 WRP-----NATAGKGVGVCNAAVYEAATKVEDGEDEPVAEDRYALGAGR--VLGAG 61  
Db 37 wrksrsvvnaavdsgrgv---vkaataavve---kpte-----rcrfeylsqk 82  
QY 62 PAPLGATALAGVNVFAVYSGATTAALCLTFPEDKADAVTEPRLDPLMNRGNWHPF 121  
Db 83 plpfgatadggnvntfavnfnataatcltllsdiprekvtteqlfdlplankqgdwhvf 142  
QY 122 IEGELHNNMLGYRFDTGFAPHCGRHLYDSNVVYDPAKAVISGEYVARGNNCPPOMA 181  
Db 143 lkgdfemmlgyrkfdgkfpeeghyfidsqglvldpakaivargeyvgipdedccppma 202  
QY 182 GMIRPLRYSTFDEMGDLPLATYPOKDLVITYBMHLRGFKHDSNVNENGTIGAVSKLDYK 241  
Db 203 gmvrpasdqfdwgdldplkfpqrdlvyemhvyrgfthnsaetkpgyglvveklidhkl 262  
QY 242 ELGVNCELMPCHEFNELEYSTSS-----KNFNGYSTINFPSPMTRTSGSIRKCGND 296  
Db 263 elgvncielmpchelefnelysynsvlgdyktnfwgstvnftspmgrysaglsncglq 322  
QY 297 AINERKTEFVREAHKRGIEVLDVFNHNTAEENGPILSFKGVDMNTYMLAPKGEFYNY 356  
Db 322 alnefkylvkeahkrgylevmdvfnhtegnekprlslfridmsvrltllapkgelyny 382  
QY 357 SGGCTFCNCPNVROFIYDCRLRYWTEMHVDFRFDLASIMTRGSSLMDPVNYGAPIE 416  
Db 383 sgcgntfcncnbpvreflvdclrywtemhvdgfrfdlasiltsssswavnvynsld 442  
QY 417 GDMITTPRLVPRPLIDMISNDPIILGAVKLIEMADAGLLYOGGPRPHNVMSKNGKTR 476  
Db 443 gdmittgprlvpplldmisdndpilaawdaggllyogprphngvisewngkyr 502  
QY 477 DIVROFIKTDGAFACLCGSPILYOGGRKPMHSINFCANHGFTLADLVYTKRKY 536  
Db 503 dmvrqfkygdgafagaefaelcgspilyaggrtkpmhsinfvcadngftladlvynh 562  
QY 537 MLPNGENNRGENHNLNMGCEGEFARLSVRLRKQRNRFVCLMVSGVPMFYMGDE 596  
Db 563 mlpngednkpgenhnmswncgegefaslfvklkrqmrnffclmvsqgpmlymgde 622  
QY 597 YGHTKGGNNNTYCHDSYNYFRMDKK-EOYSELHRRCCMTFRKCEGLGLEDPTAKR 655  
Db 623 yghtkgnnntychdshyvyfrwdkkesgsdflfcglmtkfrnceesglgdgfptraer 682

Query Match	Score 2919.5	DB 20	Length 931
Best Local Similarity	69.6%	Pred. No. 2.6e-250	

PH	Key	Location/Qualifiers
FT	Misc-difference	533
FT		/label= unknown
FT		/note= "encoded by TGA"
FT	Misc-difference	541
FT		/label= unknown
FT		/note= "encoded by TAA"
FT	Misc-difference	546
FT		/label= unknown
FT		/note= "encoded by TAG"
PN	WO914314-A1.	
PD		
PD	25-MAR-1999.	
XX		
PE	11-SEP-1998;	98WO-AU00743.
PR	20-MAR-1998;	98AU-0002509.
PR	12-SEP-1997;	97AU-0009108.
XX		
PA	(CSIR ) COMMONWEALTH SCI & IND RES ORG.	
PA	(GOOD-) GOODMAN FIELDER LTD.	
PA	(LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.	
PA	(AUSU ) UNIV AUSTRALIAN NAT.	
XX		
PI	Li Z, Morell M, Rahman S;	
DR	WPI, 1999-229525/19.	
DR	N-PSDB; AAX34654.	
XX		
PT	New isolated cereal plant enzyme genes used for, e.g. expression of	
PT	antisense sequences of granule bound synthase	
XX		
XX	Example 23: Page 100-102; 171pp; English.	
CC	The invention relates to a novel enzyme of starch biosynthetic pathway	
CC	in a cereal plant, where the enzyme is selected from starch branching	
CC	enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching	
CC	enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or	
CC	SBE I of rice or maize. The methods and products can be used for	
CC	targeting expression specifically to the endosperm of the seeds of cereal	
CC	plants such as wheat or barley. They can be used for the expression of	
CC	e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low	
CC	mol. wt. glutenin, grain softness protein I, bacterial isomylase,	
CC	bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They	
CC	can be used for modifying the characteristics of starch produced by a	
CC	plant. The present sequence represents the wheat DBE partial sequence.	
XX		
XX	Sequence 568 AA:	

	Query Match	69.3%	Score 2918:	DB 20:	Length 568:
	Best Local Similarity	98.7%:	Pred No.1.7e-250:		
	Matches 525:	Conservative 3:	Mismatches 4:	Indels 0:	Gaps 0:
QY	233	AVSKLDYKEKELGVNCEIEMPCHEPNELESTSSSKMKNFEGYSTINFSPMRTYSGIKN	292		
Db	1	avskldykekgvncetelmpcheheleystsskmmfvgystlnffspmtlytsaglkn	60		
QY	293	CGRDAINEFKTFVREAHKRGILEVILDVVFNNHTAEGNENGPILISFKGVDTTYMYLAPKE	352		
Db	61	cgrdainefktrvreaahkrglevildvfvnhtaegnengpilisfcgvdttymlapke	120		
QY	353	FVYNSGCGNTENCNHPVVRQPIVDCLRHWYTMKHADGPRFDLASIMTKSSLMDEVNYNG	412		
Db	121	fyngsgcgtlncnhpvrvgqfivdclrywmemhvdgtrfdlasimtkrgsslmdevnyng	180		
QY	413	APIEDMTTGTPLVTPPLIDMISNDPLIGVKYLLAEAMDAGGLQVGGEPFMHWNSEKN	472		
Db	181	apiedmmtctgplvtpplidmismdplligvkyllaeadagglqvvgqfpmhwnsekn	240		
QY	473	GKRIIDYKQPIKGTDFVAGGFAECLCGSPHLYQAGGRPMWHSINFCVCAHDEFTLADLVNY	532		

Db	241	gkyrdlrvqfktkctgdgfaagfaeclcgspnhlyqegarkpwshlrfvcabhdgfcflgdvlvy	300
Qy	533	NKKYLPNGENNRRDGEENHNLSWMCNCEGEFAPLRSVKRLRRKRMKNFEVCLMWSOGVPMFY	592
Db	301	nnkynlpngennrdgenhnlswmcneegefafarlsvckrlrkqrntffvcclmwsagvpmfy	360
Qy	593	MGDEGHKTKGGNNNTYCHDSYVNVFRFMDKKEOYSSELRRPCCLTMTFRKBEEGSLDEPFT	652
Db	361	mgdeghktkggnnntychdsyvvnfrfmdkkfoysseelrrpccltmtfrkbeegsldepfpt	420
Qy	653	AKRLQMHGHQKQDMSNSRFFVAFSSKDEKOGETVYAFPTSHLPAVELPEBAGRMRWP	712
Db	421	aerlqmhghqkqpdmsensrfvafsskdekergetvyaftshlpavelpebagrmrwp	480
Qy	713	VVDYGRKAPRYVFLTDLPDLRALTTHQGSHELYSLYPMLSYSSVYLALRPDV	764
Db	481	vvdygkpaipyfllcdlfpdrclrlthgshlnsllypmlyssvylalrpdv	532

RESULT	6
AAW36602	ID AAW36602 standard; Protein; 606 AA.
XX	
XX	
AC	AAW36602;
XX	
DT	08-MAY-1998 (first entry)
XX	
DE	S. tuberosum debranching enzyme.
KM	Debranching enzyme; PDE; potato; transgenic plant; starch; industry; food manufacture.
XX	
OS	Solanum tuberosum.
FH	Key
FT	Protein
FT	Location/Qualifiers 1..606 /note= "partial protein sequence"
XX	
PN	DEI9618125-A1.
XX	
PD	13-NOV-1997.
XX	
PF	06-MAY-1996; 96DE-1018125.
PR	06-MAY-1996; 96DE-1018125.
XX	
PA	(PLAN-) PLANTEC BIOECONOLOGIE GMBH.
XX	
DR	WPI: 1997-551167/51.
DR	N-PDB: AAT96754.
XX	
PT	DNA encoding potato debranching enzyme - useful for producing recombinant enzyme or transgenic plants for use in food production or in industry
PS	Claim 1; Page 18-19; 20pp; German.
XX	
CC	This sequence represents a novel potato debranching enzyme (PDE).
CC	The pDE gene can be used in the production of transgenic plants especially starch storing plants e.g. Cereal or potato plants. Other nucleic acid molecules could be designed which specifically hybridise to the pDE transcript and inhibit its translation. Vectors containing this gene can be linked in sense orientation with regulatory elements that ensure its transcription and translation in prokaryotic and eukaryotic cells. The starch produced from such transgenic plants could be used in the manufacture of foods or industrial products.
CC	
XX	
XX	
Sequence	606 AA;
Query Match	62.0%; Score 2612; DB 18; Length 606;
Best Local Similarity	76.6%; Pred. No. 3e-223;
Matches 458; Conservative 62; Mismatches 70; Indels 8; Gaps 3	

```

QY 171 ARG--NNCPROMAGMIRPLEYSTFDMEGDLPLARKYRQKDLVYTEMHNRGETKDDSSNVHPG 228
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| :|||
Db 3 atgpeddcwpmagmwpasdqfdwgdllkfpqdlvlyemlvrgflhhesetkypg 62
QY 229 TFGAVSKLDLYELKELVNCIELPCHFEFNELEKSTSS-----KMNWGSTINFESPMPT 283
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 63 tlygvveklidhkeijynclmepcheefneleyssynsvgdtkfnfvgystvnffspmg 122
QY 284 RYTSGSIKNGRDAINEFKTFVREARKRGIEVLDDVFNHTAEGNENGPILLSFGVDNPT 343
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 123 ryssaglsncjgalnefkylyvkeahkrgylevldmvdvfnhtlaegenppllsffgldnsv 182
QY 344 YVYLAKGGEFYNTSGGNTFNCNHPVYRQFTVDCLRVWTEMHVNDGRRFLASINTGSS 403
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 183 fclapkgelynysgcgnlfncompivrtqflvdclrywtemhvdgfrfdaaslltrss 242
QY 404 LMDPVNVGAPRIGDMITTTGTPPLVTPPLIDMISNDPILGKYLIAEAMAGLQVOCFP 463
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 243 swnavnvynsldgdmlltcgpltsppllmldmndpilsyvkllaaewdcggllyqvgmfp 302
QY 464 HNNWSENNKGYRDIYRQFIKGTDFAGFAECLCGSPHLYOAGRRKPMHSINFVCAHDG 523
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 303 hmgivsewngkyrdmvrqfllkgtgdfsgafaeclcgspnlyqkggrkpmwsinfvcahdg 362
QY 524 FTLADLVYTKKATYLPNGENNRDGENHNLNMGCEGEFARLSYKRLRKQMRNFCPLM 583
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 363 fltadlvtyunkhnlangednkhgenhmsnwcgegefasllyvklrlrkqmrnfflclm 422
QY 584 VSGGVMEFWYDGEYGHKGGNNNTYCHDSYVNYFRMDKK-EOYSELARPCCLMTKFRKEC 642
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 423 vsqgvpmlymgdeyghksgnnlcychnylnyfrwckkdeessdflrfcgllmktfrhec 482
QY 643 EGTGLEDEPTAKRLOMIGHOPKRPMSNSRFVAFSMKDEROGELIVAFNTSHLPAAVEL 702
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 483 esglgdfpfaerlqwhgncprtpwsetstfvaltlvdkvkgelylafnashlpvtclt 542
QY 703 PERAGRRRWEPVDTGKAPAYDFLTDLPRALTIHQFSHFLYSNLPMLSYSSVLYL 760
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 543 pekgyrwyqpfvdtgkppafdltdvperetaakqyshfidangymlyssylll 600
RESULT 7
AA04382
ID AAY04382 standard; Protein: 429 AA.
XX
AC AAY04382;
XX
DT 06-JUL-1999 (first entry)
XX
DE Potato isoamylase type DBE genomic clone protein SEQ ID NO:15.
XX
KM Potato; isoamylase debranching enzyme; DBE; starch.
XX
OS Solanum tuberosum.
XX
PN WO9912950-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-GB02665.
XX
PR 06-SEP-1997; 97GB-0018863.
XX
PA (NAT) NAT STARCH & CHEM INVESTMENT HOLDING COR.
XX
PI Jobling SA, Schwall GP, Westcott RJ;
XX
DR WPI: 1999-229220/19.
XX
DR N-PSDB: AAX33162.
XX
PT New isolated potato isoamylase-type debranching enzyme gene
XX

```

```

PS Claim 6; Fig 8; 72pp; English.
XX
CC The present invention describes a novel nucleic acid sequence which is
CC obtainable from potato plants and carries at least a portion of an
CC isoamylase-type debranching enzyme (DBE) gene. Constructs containing the
CC nucleic acid sequence can be used to alter the starch properties of
CC plants such as potato, sweet potato, maize, wheat, barley, oat, cassava,
CC pea or rice. The starch can have increased branching and/or shorter
CC chain length, reduced peak viscosity, higher setback viscosity or
CC increased viscosity onset temperature. By using an antisense sequence
CC with greater homology to the native gene, greater inhibition can be
CC achieved. The present sequence represents a protein from an isoamylase
CC type DBE genomic clone.
XX
SQ Sequence 429 AA:
XX
Query Match 44.7%; Score 1881; DB 20; Length 429;
Best Local Similarity 77.4%; Pred. No. 1,8e-158;
Matches 328; Conservative 44; Mismatches 50; Indels 2; Gaps 2;
QY 338 GVDNTTYMLAPKGEFYNSGCGTFCNCHPVRQFTVDCLRVWTEMHVNDGRRFLASI 397
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 1 gldnsvfytlapkgelynysgcgnlfncompivrtqflvdclrywtemhvdgfrfdaasll 60
QY 398 MTRGSSLMDEPVNVGAPRIGDMITTTGTPPLVTPPLIDMISNDPILGKYLIAEAMAGL 457
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 61 ltrssswnavnvynsldgdmlltcgpltsppllmldmndpilsyvkllaaewdcgglly 120
QY 458 OVQGFPMNWNSENNKGYRDIYRQFIKGTDFAGFAECLCGSPHLYOAGRRKPMHSINF 517
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 121 qvgmflmvglysewn-eytdmvrqflogtqdfsgafaeclcgspnlyqkggrkpmwsinf 179
QY 518 VCAHDGFTLADLVYTKKATYLPNGENNRDGENHNLNMGCEGEFARLSYKRLRKQMRN 577
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 180 vcahdgftladvlyunkhnlangednkhgenhmsnwcgegefasllyvklrlrkqmrn 239
QY 578 FTVCLAMYSQGYPMYMGDEYGHKGGNNNTYCHDSYVNYFRMDKK-EOYSELARPCCLMT 636
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 240 flflmwsqgvpmlymgdeyghksgnnlcychnylnyfrwckkdeessdflrfcgllm 299
QY 637 KFRKEGEGLEDEPTAKRLOMIGHOPKRPMSNSRFVAFSMKDEROGELIVAFNTSHL 696
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 300 kfineseslgldgfpfaerlqwhgncprtpwsetstfvaltlvdkvkgelylafnashl 359
QY 697 PAVVELPERAGRRRWEPVDTGKAPAYDFLTDLPRALTIHQFSHFLYSNLPMLSYSSV 756
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 360 pvtcltldprgyrwyqpfvdtgkppafdltdvperetaakqyshfidangymlyssyl 419
QY 757 ILVL 760
   ||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||
Db 420 llll 423
RESULT 8
AA00870
ID AAY00870 standard; Protein: 766 AA.
XX
AC AAY00870;
XX
DT 21-MAY-1999 (first entry)
XX
DE S. tuberosum isoamylase clone 9 protein sequence.
XX
KM Isoamylase clone; potato; branched polysaccharide synthesis; amylopectin;
XX starch modification; debranching activity.
XX
OS Solanum tuberosum.
XX
PN WO9906575-A1.
XX
PD 11-FEB-1999.
XX

```

PF 30-JUL-1998; 98MO-GB02280.  
XX 31-JUL-1997; 97GB-0016185.  
XX (PLAN-) PLANT BIOSCIENCE LTD.  
PA Bustos Guillen R, Edwards EA, Martin CR, Smith AM;  
PI WPI: 1999-153803/13.  
DR N-PSDB; AAX27062.  
XX  
XX New nucleic acid from potato encoding starch debranching enzyme with  
PT isomylase activity - and related vectors, transformed cells and  
PT plants, proteins and antibodies, used to generate starch with  
PT altered properties, for use in foods, as thickeners etc.  
PS Claim 2: Flg 6; 78pp; English.  
XX  
XX This sequence is a Solanum tuberosum isomylase of the invention.  
CC Fragments of the isomylase coding sequence are used to identify and  
CC clone isomylases from other plant species, by standard hybridisation  
CC or amplification methods. Expression of the isomylase DNA in host cells  
CC is used to produce the isomylase which is used: (1) to raise antibodies  
CC (Ab) for identification, isolation and localisation of isomylases; and  
CC (11) for synthesis of branched polysaccharides. The DNA, proteins  
CC and Ab (or related peptides) are used to alter the quality and  
CC quantity of polysaccharides in a host cells, specifically to alter  
CC branching in amylopectin. The modified starches produced are useful in  
CC human or animal foods; as biodegradable plastic; as food or paint  
CC thickener; in starch-coated films, paper and textiles; in milking  
CC explosives; in pharmaceuticals and glues. Manipulation of debranching  
CC activity in a plant allows control of starch properties, e.g. increased  
CC gel strength; formation of paste rather than gel, changed physical  
CC characteristics etc.  
XX  
XX Sequence 766 AA;  
SQ

Query Match 37.4%; Score 1577; DB 20; Length 766;  
Best Local Similarity 44.0%; Pred. No. 4.9e-131;  
Matches 329; Conservative 111; Mismatches 235; Indels 72; Gaps 19;

QY 30 AATKEDGCEDEPVAEDRYALGACRVLAGMPAPUGATALGCVNFVSGGATTAALC 89  
DB 70 aaakqgeaa---pqmltdif---sfkvspglahpdyvsetesgfalfsqhasavtlc 122  
QY 90 LETPEDLKADRYTEVPIDPLMNRGTGNVHVFIEGEL--HNLXGVRFGTAPHCCHTL 147  
DB 123 lllpsvhdgml--elaldpqkntqgdllwhicik-elpqgsvlygyrldqprnwheghrf 179  
QY 148 DVSNVVDVPYAKAVISRGEGYVARGNMCWPMAGMIRLPYSTFDEGDLPL-RYPQKDL 206  
DB 180 ddslllvqpyeklllgtrvfg---desnkmercflytdfnsjpfdwgenykipnlpbekl 236  
QY 207 VYEMHLRGFTKHDSNNEHP--GTFIGAVSKLDVLEKLGVCNCLIMPCHFEFNELEY--- 261  
DB 237 viyemvrafaadetsalddqgrsygljlekipnllelynavellpyfeddelelqrr 296  
QY 262 -STSSKNNFNGCYSTINFFSPWTRTSGCIKNCGRDAIN---EFTFVFAEARKRGIEVL 317  
DB 297 nuprdlmhlnatwycstlnftapmsryas-----cgggprvasawefkemvkalhgaaglevll 351  
QY 318 DVFNHTAEGNENGP-IISFKGVDMTTYM--LAPKGEFYNSSGCGTFCNHPVYRORI 374  
DB 352 dvynhttheadenpytsfsgldnkkyymvwlinnagllfatgcnltfnchpvcymel 411  
QY 375 VDCLRVWTEMHVDFRFRDASIMTRGSSLLMPDVNVYAGPIEGDMITTGTPVTPPLIDM 434  
DB 412 leslrhwteyhdvdfirdlaavlcry-----tdgtfplnapllyvka 452  
QY 435 ISNDPILGCVTLIAAMAGGLYOGGFRPHWNVSEMNCKRYDIROAFKRGIDGFRGGA 494  
DB 453 lskdsvlstrckllaeprwcdggllylvgkfpnwdrwaewngkyrdldlrrllkdgamkgnfa 512

QY 495 ECLCGSPHLYOAGRRPWHNSINFVCAHDGFTLADLYTYNKKYMLPNCENNRDENNHLSW 554  
DB 513 trlsgsadlyrvnkkrkpyhvnftlahdgtfilydvsynknhdangegngndnfnfsw 572  
QY 555 NCGEDEGFARLSVKRLRKRMRFVCLAMYSQGVPMFMDDEGHTKGGNNITYCHDSYV 614  
DB 573 ncgiletsdanlnalstrsmknfhlaIwvsgtupmmlmgdeyghtrlygnnsyghdtaI 632  
QY 615 NYFRMCKKEQYSELH--RCCCLMTKFRKCEGGLGLEDPTAKRLQWGHGQKPPM-SENS 672  
DB 633 nftqwgglearkndhfrfsskmlkfrlshnvltkenfiekndltwl-----edwvnees 687  
QY 673 RFAVFSMKDEROGEIYVAFNTSHLPVAVELPE-RAGRRPEVYDTGKRPAPYDFTDLPD 731  
DB 688 rfiafmlhdngggdilylafnabhfslkltalpspprnmsyrvvtdnlkspddftvegsvg 747  
QY 732 RALTIHQSFHLXSNLXPMLSYSVIL 758  
DB 748 -----lsktydvepyaalll 762

RESULT 9  
ID AAY34991 standard; Protein; 666 AA.  
XX AAY34991;  
AC  
XX  
DT 13-SEP-1999 (first entry)  
XX  
DE C. pneumoniae protein involved in intermediate metabolism.  
XX  
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX  
XX Chlamydia pneumoniae.  
OS  
XX  
PN WO9927105-A2.  
XX  
PD 03-JUN-1999.  
XX  
XX 20-NOV-1998; 98MO-IB01890.  
PF 04-NOV-1998; 98US-0107078.  
PR 21-NOV-1997; 97ER-0014673.  
XX  
XX (GEST ) GENSET.  
PA  
PI  
PI Griffiths R;  
DR WPI: 1999-357842/30.  
XX  
XX Genome sequence of Chlamydia pneumoniae  
PT Page 898-899; Disclosure; 1912pp; English.  
XX  
XX AAY34584-Y35879 represent the proteins encoded by all the open reading  
CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.  
CC  
XX  
SQ Sequence 666 AA;

Query Match 31.0%; Score 1305.5; DB 20; Length 666;







XX Starch conversion processes - using isoamylase(s), particularly  
 PT obtained using novel DNA cloned from Rhodothermus or Sulfolobus  
 PT strains

PS Claim 23; Pages 67-69; 76pp; English.

XX The sequence is that of an isoamylase which can be used as part of a  
 CC method of starch conversion for the production of high fructose  
 CC syrup or fat replacers. The use of an isoamylase in starch conversion  
 CC results in a reduced formation of the undesired trisaccharide panose  
 CC to increase the saccharification yield.

XX Sequence 726 AA:

Query Match 28.2%; Score 1187; DB 20; Length 726;

Best Local Similarity 38.9%; Pred. No. 2.2e-96;

Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

QY 57 VLACMPAPLACATACAGVNFAYVSGATAAALCLFT-PEDLKADRYTEVPPLDPLMNRG 115  
 DB 14 VWPGRPYPLGATGAGVNFAYVSGATAAALCLFT-PEDLKADRYTEVPPLDPLMNRG 115  
 QY 116 NVMHVFIEGELNMLGYRFDGTFAPHCGLDVSNNVVDPAKAVISRGEGYPARGNN 175  
 DB 68 PLWHVYLPGLIPGQLYGYRYPYRPEEGHFINPKVILDPYAKAI-----GRPLR--- 118  
 QY 176 CPMQW-----AGMIP-----PYS-----TFDMEGDLPLRYPOKDLVIREMH 212  
 DB 119 -WDSLIFYKIGDPAGDLISEEDSAPYAPLGVAVEGCEFGWGDPRPLMEDLLYETH 177  
 QY 213 LRQFTRHDSNVNHP-----GTFIGAVSK--LDYLKELGVNCELMPCHEFNELEYST 263  
 DB 178 VKGLTK-----LHPEVPEPLRGTYLGLTCEPYLEHLKQLYGVTFIGLIPYNAKHVHDHLV 231  
 QY 264 SSSKAMFWGISTINFSFPMRFR--TSGCIKNCGRDAINEFTFVEAKRGCEVILDVVEN 322  
 DB 232 ERLIRYWGYNPLCYLTPAPEYALNPL-----SAVREKMMVVALHAAGFEVILVDVYV 286  
 QY 323 HTAEGNNGNGLISFKGVNDNTYYMALAPKGEFY--NYSGCCNTFNCNHPVYRQFIVDCLRY 380  
 DB 287 HTEGGVLPGLTIFRGDNTAYKADPNRFLVDYGTGNTLDVGNPYVQLMDLRY 346  
 QY 381 WTEMHVDGFRDLASTMTKSSLSMDPVNNGADIEGDMITGTPLVTPPLIDMISNDPT 440  
 DB 347 WTEMHVDGFRDLASTMTKSSLSMDPVNNGADIEGDMITGTPLVTPPLIDMISNDPT 440  
 QY 441 LGGVKILAEAMDAG-GLYVGOPPHMNVWSEMGKYRDIYRQFIRKGTDFAGFEACLCG 499  
 DB 387 LAGVKILAEAMDAG-GLYVGOPPHMNVWSEMGKYRDIYRQFIRKGTDFAGFEACLCG 499  
 QY 500 SPLUYQAGKRRPMHSINFCVADHDTLADLVYNNKKNLPGNENNRDGENHLSMNGCEE 559  
 DB 445 SSDLTSGRPTFASINFCVADHDTLADLVYNNKKNLPGNENNRDGENHLSMNGCEE 559  
 QY 560 GEFARLSVKRLRQRMRNFVYCLWVSGVPEFYWGDEYGHGHTGNNNTYCHDSIVNFRW 619  
 DB 505 GPQDQPEVLACREALKRLSLSTLFLISGVPMILGDELSTQHGNNAYCQDNLISWYNW 564  
 QY 620 D-----KKQYSELHRCCLMTRKKECEGLGLEDFTAKRLQWQHQP-----GRP-- 666  
 DB 565 QDLTRKQFLE-----IVRQTLIRKQHSFIRRHILGLP--NGSRPTSLVAPEGIRPMR 618  
 QY 667 --DMSSESRVAFSM-----KDE-----RQGEIVAFNTSLHPLAVELPERAG--- 707  
 DB 619 HEDVC-NPELTAFGILLHGDAIGTDEHGPRFIDDTLILFMNGSEAVPVVVEVSCGK 677  
 QY 708 -RRMEPV-----VDTGKPAVYDELTDLPDRALTI 736  
 DB 678 PLWHVYLPGLIPGQLYGYRYPYRPEEGHFINPKVILDPYAKAI-----GRPLR--- 118

RESULT 13

AAV78515

ID AAV78515 standard; protein; 726 AA.

AC AAV78515;

DT 05-MAY-2000 (first entry)

DE Isoamylase amino acid sequence.

XX Isoamylase; debranching enzyme; amylopectin; thermostable; corn;  
 KW genetically engineered variant; wheat; potato; wheat; manioc; rice;  
 KM starch; sugar production; high fructose syrup; high maltose syrup;  
 KM maltodextrin.

XX Rhodothermus marinus.

XX WO200001796-A2.

XX 13-JAN-2000.

XX 02-JUL-1999; 99WO-DK00381.

XX 02-JUL-1998; 98DK-0000868.

XX (NOVO) NOVO-NORDISK AS.

XX Bisgard-Frantzen H, Svendsen A;

XX WPI; 2000-160767/14.

XX N-PSDB; AA289961.

PT Variant bacterial pullulanases and isoamylases having, e.g. increased  
 PT thermostability, used for converting starch from potatoes into high  
 PT fructose syrup

XX Claim 29; Fig 3; 116pp; English.

CC This sequence represents an isoamylase amino acid sequence. Isoamylases  
 CC are debranching enzymes which hydrolyse alpha-1,6-D-glucosidic branch  
 CC linkages in amylopectin. This sequence represents a parent enzyme from  
 CC which genetically engineered variant debranching enzymes can be created.  
 CC The genetically engineered enzyme has an improved thermostability at a pH  
 CC in the range of 4-6 compared to the parent enzyme. The modified enzyme  
 CC may alternatively or also have an increased activity towards amylopectin  
 CC and/or glycogen compared to the parent enzyme. The enzyme variants are  
 CC useful for converting starch to one or more sugars. This is useful when  
 CC starches such as corn, potato, wheat, manioc and rice starch are used as  
 CC starting materials in commercial large scale production of sugars, such  
 CC as high fructose syrup, high maltose syrup, maltodextrins, amylose, G4-G6  
 CC oligosaccharides and other carbohydrates such as fat replacers. The  
 CC thermostable debranching enzymes of the invention make it possible to  
 CC perform a liquefaction and debranching simultaneously before the  
 CC saccharification step. By debranching using the thermostable debranching  
 CC enzyme variants during liquefaction together with the action of an  
 CC alpha-amylase, the formation of panose precursors is reduced, thereby  
 CC reducing the panose content in the final product and increasing the  
 CC overall saccharification yield.

XX Sequence 726 AA;

Query Match 28.2%; Score 1187; DB 21; Length 726;

Best Local Similarity 38.9%; Pred. No. 2.2e-96;

Matches 294; Conservative 86; Mismatches 242; Indels 134; Gaps 30;

QY 57 VLACMPAPLACATACAGVNFAYVSGATAAALCLFT-PEDLKADRYTEVPPLDPLMNRG 115  
 DB 14 VWPGRPYPLGATGAGVNFAYVSGATAAALCLFT-PEDLKADRYTEVPPLDPLMNRG 115  
 QY 116 NVMHVFIEGELNMLGYRFDGTFAPHCGLDVSNNVVDPAKAVISRGEGYPARGNN 175  
 DB 68 PLWHVYLPGLIPGQLYGYRYPYRPEEGHFINPKVILDPYAKAI-----GRPLR--- 118

```

QY 176 CWPQM-----AGMIFL-----PYS-----TFDMEGDLDLRYPQKDLVYEMH 212
Dy 119 whdalfgykldpgddisfseedsapypdlgavvegcfewgddrpprlpmedtlllyeth 177
QY 213 LGFTKHDSSNVEHP-----GTFIGAVSK--LDYLKELGVNCIELMPCHEFNELEYST 263
Dy 178 vkglck-----lhpevrplrgtylgldceprlehkhqlygttqlqlprhkvhrthlv 231
QY 264 SSSKKNFNGYSTINFEPSPTRY--TSGGICNCGDAINEKTEFYREAHKRCIEVILDVEN 322
Dy 232 ergltnywgynplcyfapepeyatngpl-----savrefkmmvralhaagfeivdvvyv 286
QY 323 HTAECNENGPITSGFVNDTFTYMLAPKGEFY--NMSGGNFNCNHPVVRQITVCLRY 380
Dy 287 htgevgvlgprclstfgrldmrraykddpnprtlldvdytgcntldvgnpyvdlqllmslry 346
QY 381 WYTEMHVDGFRDLASIMRGSLLMDPVNVGAPLEGDMITITGTPPLVTPPLDIMISNDP 440
Dy 347 wtemhvdgfrldlaaalr--elydv-----dmnst-----ffgvlqddrv 386
QY 441 LGVVLLEANDAG--GLTVOGQPPHMNVSEWNGKRYDVRQFTKGTDFAGGAECLCG 499
Dy 387 lsgvllaeppwvpgpygynhfr-wq-wtemngryrdavrrfwrgdrlngefatrrfag 444
QY 500 SPHLQAGGRKPMHSINFCAHDGFTLADLVYTNKKYNLPNGEENNDGNNHNSWNGGEE 559
Dy 445 ssdlversgrtrpfaasintvlandgtlledvaytkkhaanalegndgndenydstcgye 504
QY 560 GEFARLKYRLRKROMRNFVCLMYSGVPMFMDGDEYGHTRKGNNTCHDSYVNYEFM 619
Dy 505 grldgprvlacteaalkrlstlflsdyvpmllsgdelstgrtghnmnyccqndelsvynv 564
QY 620 D---KKEQYSELHNRCCCLMTKRKCEGELGDEPFAKRLQNHQNP-----GKP-- 666
Dy 565 qldrtkgqfle---fvrgtlwfrtkhstfrtrhfltlgdp---nggrprslvapegrprmr 618
QY 667 --DMSENSRFVAFSM-----KDE-----ROGEIYAFWTSHLPAVELPERAG--- 707
Dy 619 hedeac-npeltatgllngdalqgtcdengrfrtdtcllllmgseavvvpvovscgk 677
QY 708 -RRMEPV-----VDTGKRPARYDFTLDDLPRALRTI 736
Dy 678 phwevvpvfgmveprcpcapgetls--lppvylv 711

```

RESULT 14  
AAU60903  
AAU60903 standard; Proteins: 718 AA.

AC AAU60903;  
Dy 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #21799.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX dermatological; osteoporotic; neuroprotectant.

OS Propionibacterium acnes.

XX MO200181581-A2.  
PN 01-NOV-2001.  
PD 20-APR-2001; 2001WO-US12865.  
PF 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX

```

PA (CORI-) CORIXA CORP.
XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX
DR WPI: 2001-616774/71.
XX N-PSDB; AAS59613.
PT
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 22098; 1069bp; English.
CC
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC to therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
CC Sequence 718 AA.
SQ

```

Query Match 25.2%; Score 1062.5; DB 22; Length 718;  
Best Local Similarity 35.5%; Pred. No. 2,5e-85;  
Matches 254; Conservative 102; Mismatches 285; Indels 75; Gaps 20;

```

QY 65 LGATLACGVNRAVYSGGTAALCLTFPBDLKAADRYEEVRLDPLMRTGNWNYFIEG 124
Dy 26 lgstlldggcrtfglwpraearevelavadd---gsgvnhmdtrd-----ehywtvevsg 77
QY 125 ELJHNLXGYRFDGTFRPHCGHYLDVSNVVDPRAKAVISGEGYVAPRG---NCKPQM 180
Dy 78 vgaqgrgyrvhngpwrpdrctgmrfnpkllldpyaraltagvdyhgrlmdhtcpesnyepdl 137
QY 181 ---AGMIFLPSYTFWDEGDLPL--RYROKDLVYTEMHNLGFTKHDSSNVEH--RGTIGAV 234
Dy 138 tddatsvplavvldopgprptlarrndisesvylethvaygtrllpdlvphlrgtyagla 197
QY 235 ---SKDYLKELGVNCIELMPCHEFNELEYSTSSKKNFNGYSTINFEPSPTRYTSGGICN 292
Dy 198 yvavlehlkslgtatelllpvgfvsqerflvgrglanygynltgffapaaecsvsng 257
QY 293 CGRDADINEKTEVREAHKRCIEVILDVGNHRAEGENGRPLISFKCVDDNTTYMLA-PRG 351
Dy 258 ---tqvrefkdmvtafheaglevflidvynhbggghegrplstfgrldhesyylendhr 314
QY 352 EEFYNSGCGNTFNCNHPVVRQITVCLRYWVTEMHVDGFRPDLASIMTSGSSIMDPVNY 411
Dy 315 ndydvrgcgnsvdthpveklamvldsltywvemgydtrfydaltclltdks-----h 367
QY 412 GAPLEGDMITITGTPPLVTPPLDIMISNDPILGCVKLLEAMDAAGL-YQVGQF-PHMNVMS 469
Dy 368 g-----vdqnhl fkgalve-----dprllkvkhlaepwdlpgpygygavagph---ws 412
QY 470 EWNKGYRDIIVRQIFGTGDFAGGAECLCGSHLYQAGGRKPMHSINFCAHDGFTLADL 529
Dy 413 ewndftrnyvdflwrgavrgveelatrlcsgsdpdy---gqtsavvntlandgftmrdl 468

```



d

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 31, 2002, 12:00:20 ; Search time 25.38 Seconds

(Without alignments)  
735,270 Million cell updates/sec

Title: US-09-674-817A-2

Perfect score: 4212  
Sequence: 1 SEPAPRLRRPNATAGKV.....SNLXPLSYSLVILRPDV 764

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Capext 0.5  
231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3527.5	83.7	818	2	US-08-410-784A-2
2	3527.5	83.7	818	4	US-09-346-237-11
3	3465	82.3	740	2	US-08-410-784A-4
4	2612	62.0	606	4	US-09-187-124-2
5	1229.5	29.2	713	4	US-09-346-237-9
6	1211	28.8	718	4	US-09-346-237-10
7	1187	28.2	726	3	US-09-129-075-4
8	1187	28.2	726	4	US-09-346-237-3
9	1187	28.2	726	4	US-09-346-237-13
10	825	19.6	774	4	US-09-346-237-8
11	810	19.2	777	1	US-08-476-519-2
12	810	19.2	777	5	PCT-US95-09323-2
13	807	19.2	746	1	US-08-476-519-11
14	807	19.2	746	5	PCT-US95-09323-11
15	803.5	19.1	776	4	US-09-346-237-7
16	802.5	19.1	776	4	US-09-346-237-4
17	782.5	18.6	750	6	5457037-3
18	781.5	18.6	751	6	5457037-5
19	685.5	16.3	772	2	US-08-410-784A-5
20	552.5	13.1	915	4	US-09-346-237-2
21	544.5	12.9	820	4	US-09-313-677-21
22	544.5	12.9	926	4	US-09-313-677-2
23	544.5	12.9	928	4	US-09-514-599-4
24	544.5	12.9	933	4	US-09-313-677-19
25	544.5	12.9	967	4	US-09-313-677-17
26	516	12.3	921	4	US-09-514-599-2
27	497.5	11.8	928	1	US-08-474-140-11

28	497.5	11.8	928	1	US-08-477-630-11	Sequence 11, Appl
29	497.5	11.8	928	1	US-08-472-293-11	Sequence 11, Appl
30	497.5	11.8	928	1	US-08-474-545-11	Sequence 11, Appl
31	497.5	11.8	928	2	US-08-478-341-11	Sequence 11, Appl
32	497.5	11.8	928	3	US-08-996-733-11	Sequence 6, Appl1
33	478	11.3	829	4	US-09-514-599-6	Sequence 1, Appl1
34	478	11.3	862	4	US-09-346-237-1	Sequence 4, Appl1
35	449.5	10.7	893	4	US-09-514-302-4	Sequence 1, Appl1
36	449.5	10.7	1338	4	US-09-514-302-2	Sequence 18, Appl
37	344	8.2	964	3	US-08-860-339-18	Sequence 15, Appl
38	339	8.0	559	4	US-09-242-690A-15	Sequence 4, Appl1
39	310	7.4	597	1	US-08-399-646-4	Sequence 4, Appl1
40	310	7.4	597	1	US-08-607-321-4	Sequence 4, Appl1
41	310	7.4	597	2	US-08-605-501-4	Sequence 4, Appl1
42	310	7.4	597	2	US-08-605-501-4	Sequence 4, Appl1
43	310	7.4	598	1	US-08-399-646-14	Sequence 14, Appl
44	310	7.4	598	1	US-08-607-321-14	Sequence 14, Appl
45	310	7.4	598	2	US-08-961-240-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-08-410-784A-2  
; Sequence 2, Application US/08410784A  
; Patent No. 5912413  
; GENERAL INFORMATION:  
; APPLICANT: MYERS, ALAN M.  
; TITLE OF INVENTION: ISOLATION OF SUL A STARCH DEBRANCHING  
; TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Welington, Schurglin, Gagnebin and Hayes LLP  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/410,784A  
; FILING DATE: 24-MAR-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heine, Ph.D., Holliday C  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: ISG-002XX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-2290  
; TELEFAX: 617-451-0313  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 818 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; ORIGINAL SOURCE:  
; US-08-410-784A-2

Query Match	83.7%;	Score 3527.5;	DB 2;	Length 818;
Best Local Similarity	83.0%;	Pred. NO. 0;		
Matches 649;	Conservative 44;	Mismatches 70;	Indels 19;	Gaps 6;

OY	1	SGAPRRLR-----RM-----RPNATGKGVGEV-----CAAVYEATATVEDGEDEP	44
Db	38	SSRPPLLAIVAGCKRRKACVGRKRPN-VAGLGIGRLSLHAAAARPAVEVQAEEDDDDEEV	96
OY	45	AEDRYALGACACRYLAGMPADLGATATAGVNFAYVSGGATAALCLPTEDLKADRTEE	104
Db	97	AEEERFALGACACRYLAGMPADLGATATALRGVNFAYVSSGASASTSLFAPDOKADRTEE	156
OY	105	VPLDPLMNRFGNWHVTEG-ELHNMLYGRFGGTAPRPHCGHLLDYSNNVYDPAAYVS	163
Db	157	VPLDPLMNRFGNWHVTEG-ELHNMLYGRFGGTAPRPHCGHLLDYSNNVYDPAAYVS	216
OY	164	RGEYGVARGNCGCPQQAAGMPLPYSTFDEGDDPLRYPKDLYIEMHLRGFTKDDSN	223
Db	217	RGEYGVARGNCGCPQQAAGMPLPYSTFDEGDDPLRYPKDLYIEMHLRGFTKDDSN	276
OY	224	VEHGTPEIGAVSKLDYIKELGVNCEILMPCHENELEYSTSSKMNFWGJSTINFSPMT	283
Db	277	TKHGTGYIGAVSKLDHKLKEGVNCEILMPCHENELEYSTSSKMNFWGJSTINFSPMA	336
OY	284	RYTSGGIRKNGCRDILNEFKTFVRAHNRKGLIYLDVVFENNTAGNENGRPLSTKGYDNT	343
Db	337	RYSSSGIRDSGCCGAINFKAFFVRAHNRKGLIYLDVVFENNTAGNENGRPLSTKGYDNT	396
OY	344	YYMLAPKGEFFNYSGCCTFNCFNHPVVRQFIVDCLRYVWTEHMHVDGFRPLASIMTRGSS	403
Db	397	YYMLAPKGEFFNYSGCCTFNCFNHPVVRQFIVDCLRYVWTEHMHVDGFRPLASIMTRGSS	456
OY	404	LMDEVNYYGABIEGDMITTTGTPLVTPRLDMISNDPILGSKVLLAEAMDAAGLYVGQRP	463
Db	457	LMDEVNYYGABIEGDMITTTGTPLVTPRLDMISNDPILGSKVLLAEAMDAAGLYVGQRP	516
OY	464	HMNWSSENGKYYRDIVAQFIKGTGTFAGFAGFACLCSSPHLYXAGGRKPMHSINFCVAHGS	523
Db	517	HMNWSSENGKYYRDIVAQFIKGTGTFAGFAGFACLCSSPHLYXAGGRKPMHSINFCVAHGS	576
OY	524	FTLADLVYNNKYYMLPNGENNRDDEHNHLSWNGCEEESEFARLSYKRLRKQRMNFFVCLM	583
Db	577	FTLADLVYNNKYYMLPNGENNRDDEHNHLSWNGCEEESEFARLSYKRLRKQRMNFFVCLM	636
OY	584	VSQGVPMFYWGEDYGHTKGKNNTTYCHDSYVNFRRMDK-EQYSELHRECCMLTKERKEC	642
Db	637	VSQGVPMFYWGEDYGHTKGKNNTTYCHDSYVNFRRMDK-EQYSELHRECCMLTKERKEC	696
OY	643	EGCLLEDPRTAKKRLQMHGHOGKGRDMSSENSFVAFSKDKEROGTYIAFNTSHLPANVEL	702
Db	697	EGCLLEDPRTAKKRLQMHGHOGKGRDMSSENSFVAFSKDKEROGTYIAFNTSHLPANVEL	756
OY	703	PERAGRRNEPVPYDCKGAPYDFLTDLDLRALTIHQESHPLSYNLYPMLSYSSVYLTLRP	762
Db	757	PERAGRRNEPVPYDCKGAPYDFLTDLDLRALTIHQESHPLSYNLYPMLSYSSVYLTLRP	816
OY	763	DV 764	
Db	817	DV 818	
RESULT 2			
US-09-346-237-11 Application US/09346237A			
: Sequence 11, Application US/09346237A			
: Patent No. 6265197			
: GENERAL INFORMATION:			
: APPLICANT: Bissard-Frantzen, Henrik			
: APPLICANT: Svendsen, Allan			
: TITLE OF INVENTION: Starch Debranching Enzymes			
: FILE REFERENCE: 5629-200-US			
: CURRENT APPLICATION NUMBER: US/09/346, 237A			

```

: CURRENT FILING DATE: 1999-07-01
: EARLIER APPLICATION NUMBER: PA 1998 00868
: EARLIER FILING DATE: 1998-07-02
: EARLIER APPLICATION NUMBER: 60/094.353
: EARLIER FILING DATE: 1998-07-28
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 818
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: PEPTIDE
: LOCATION: (1)..(818)
: OTHER INFORMATION: Isoamylase
US-09-346-237-11

```

Query Match	83.78;	Score 3527.5;	DB 4;	length 818;
Best Local Similarity	83.08;	Pred. No. 0;		
Matches 649;	Conservative 44;	Mismatches 70;	Indels 19;	Gaps 6;

QY	1	SGPARLR-----RW-----	RENAATGAGVGV-----	CAAVNEAATVDEBEGEDEYV	44
Db	38	SSPRLLAVPAGRMVAGVGRGRN-	VAGUGRGLSLHAAAARVAEVAQ	EDDDDDDEY	96
QY	45	AEDRYALGACRYLGLMPAPL	GCATYLAGGVNFAVYSGCAT	ATAALCTPREDLKADRYTEE	104
Db	97	AEEERLALGACRYLGLMPAPL	GCATYLAGGVNFAVYSSGAS	ASALSLFLAPDALKADRYTEE	156
QY	105	VLDDLMNRGTGMVHVPLEG-	ELHMLKVGREDGFAPACGY	LQVSNVVDYAAVYS	163
Db	157	VLDDLMNRGTGMVHVFLH	DELDHMLCQYRDGFAP	ERGOYDVSNNVDDYAAVYS	216
QY	164	RGEYVPARGNCCPMQAMG	MIPLPYSTEDMEDDLRY	YPOKDLVYIEMLRGFTKHDSSN	223
Db	217	RGEYVPARGGSCWQMAG	MIPLPNKPEMODDLP	GYQKDLVYIEMLRGTKINSXK	276
QY	224	VEHPGTFGAVSKLDYL	LAELGVNCTELMCPHE	NELEYSTSSSKNFMWYSTINFPSPMT	283
Db	277	TKHPGTFGAVSKLDHL	KLGVNCTELMCPHE	NELEYSTSSSKNFMWYSTINFPSPMA	336
QY	284	RTSGGIRKCGDALIN	PERTVYREAHKRIEY	LDVYENHTABGNNGPILSRKGVDNIT	343
Db	337	RYSSGIRBSGGAIN	EFKAFVREAHKRIEY	IMDVFMHTAEGNKRGILSRGIDNSR	396
QY	344	YMLPKRGFEYVYSG	GNFENCNHPVRG	FOYDCLRYVTENHVDREFRDLASIMRGSS	403
Db	397	YMLPKRGFEYVYSG	GNFENCNHPVRG	FOYDCLRYVTENHVDREFRDLASILRGCS	456
QY	404	LMDPVNVYAPIT	EGDMITTGTPLYVR	PLDLMSNDPILGCVKLIAEAM	DAGSLYQVGOFP 463
Db	457	LMDPVNVYSGPME	DMITTGTPLYVR	PLDLMSNDPILGCVKLIAEAM	DAGSLYQVGOFP 516
QY	464	HMNVSENNGKTRD	YRQPIKGTDFAGF	ABCLGSPHLYQAGKRP	PMHSINFCVACHD 523
Db	517	HMNVSENNGKTRD	YRQPIKGTDFAGF	ABCLGSPHLYQAGKRP	PMHSINFCVACHD 576
QY	524	FTLADLVYNNKRYN	LPNGNRRDGENH	NLSNMCGEGEPARLSV	KRLKROMNPFVCLM 583
Db	577	FTLADLVYNNKRYN	LPNGNRRDGENH	NLSNMCGEGEPARLSV	KRLKROMNPFVCLM 636
QY	584	VSGQVPMFAGDEY	GHTKCGNNNTYCH	DSYVYFMRDK-K-EQYSEL	HRCCCLMTTFRKED 642
Db	637	VSGQVPMFAGDEY	GHTKCGNNNTYCH	DSYVYFMRDK-K-EQYSEL	HRCCCLMTTFRKED 696
QY	643	ESLGEDFPTASRL	KWHGQPKPDM	SEASREFAFMKDETG	KEIYVAENTSHLPVYGL 702
Db	697	ESLGEDFPTASRL	KWHGQPKPDM	SEASREFAFMKDETG	KEIYVAENTSHLPVYGL 756
QY	703	PSRAGRMPEVVD	GKPARYDFLTDL	PLRALITHQSFHVS	NLYPMLSYSSVILVLR 762
Db	757	PSRAGRMPEVVD	GKPARYDFLTDL	PLRALITHQSFHVS	NLYPMLSYSSVILVLR 816



OY 763 DV 764  
 DB 817 DV 818

# RESULT 3

US-08-410-784A-4  
 Sequence 4, Application US/08410784A  
 Patent No. 5912413  
 GENERAL INFORMATION:  
 APPLICANT: MYERS, ALAN M.  
 APPLICANT: JAMES, MARTHA G.  
 TITLE OF INVENTION: ISOLATION OF SUI, A STARCH DEBRANCHING  
 TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE  
 TITLE OF INVENTION: SUGAR 1  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP  
 STREET: Ten Post Office Square  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/410,784A  
 FILING DATE: 24-MAR-1995  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Heine, Ph.D., Holliday C  
 REGISTRATION NUMBER: 34,346  
 REFERENCE/DOCKET NUMBER: ISU-0022XX  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-542-2290  
 TELEFAX: 617-451-0313  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 740 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 US-08-410-784A-4

Query Match 82.3%; Score 3465; DB 2; Length 740;  
 Best Local Similarity 85.3%; Pred. No. 0;

Matches 631; Conservative 43; Mismatches 62; Indels 4; Gaps 3;

OY 27 VVEAATKVEDEGEDEDEPAEDRYALGACRVLAGMPAPLGTALAGVNFAYVSGATAA 86  
 DB 3 VAEAVQAEEEDDDDDDEVAERFALGACRVLAGMPAPLGTALRGVNFAYVSSGASAA 62  
 OY 87 ALCLFTPEDLKADRVTEEVPLDPLMNRGTGNVHYFIEG-ELHNMUYGTRDGTAPHCGH 145  
 DB 63 SLSTFAGDGLKADRVTEEVPLDPLMNRGTGNVHYFIEGDELHGMCGYRFGVRAPEBQ 122  
 OY 146 YLDVSNVYVPYAKAVYISRGEGVPARGNMCPOMAGMIPPLYSTFDEGDLPLRYPOKD 205  
 DB 123 YLDVSNVYVPYAKAVYISRGEGVPARGNMCPOM--MIPPLYNKFDQGLPLRYPOKD 180

OY 206 LVIYEMHLRGFTKHDSSNVEHDEPTFICAVSKLDYKELGVNCTELMPCHENFLEYSTSS 265  
 DB 181 LVIYEMHLRGFTKHDSSNVEHDEPTFICAVSKLDYKELGVNCTELMPCHENFLEYSTSS 240  
 OY 266 SKMNFVGSTINFEFPMRTYSGGJIKNGCRDAIINEKTVREAHKRGIEVIIDVFNHNA 325  
 DB 241 SKMNFVGSTINFEFPMRTYSGGJIKNGCRDAIINEKTVREAHKRGIEVIIDVFNHNA 300  
 OY 326 EGNENGPILSFQVNDTYYMLAPKGEFYNVSGCGTFCNPNPVYQFIVDCIARYVTEH 385  
 DB 301 EGNENGPILSFQVNDTYYMLAPKGEFYNVSGCGTFCNPNPVYQFIVDCIARYVTEH 360  
 OY 386 HYDGFREFDLASITRGSSLMDEPVNYGADIEGDMITGTPVLPPLIDMISNDPIIGYK 445  
 DB 361 HYDGFREFDLASITRGSSLMDEPVNYGADIEGDMITGTPVLPPLIDMISNDPIIGYK 420  
 OY 446 LIAEAMDAGLQVOQGFPMWNVSEKNGYRQIVRQFICTGTGAFAGFACLCGSPHLQ 505  
 DB 421 LIAEAMDAGLQVOQGFPMWNVSEKNGYRQIVRQFICTGTGAFAGFACLCGSPHLQ 480  
 OY 506 AGGRKPMHSINFCVADHGTFLADLVYNNKYMLPNCNNRDEBNHLSNNGCEGEPARL 565  
 DB 481 AGGRKPMHSINFCVADHGTFLADLVYNNKYMLPNCNNRDEBNHLSNNGCEGEPARL 540  
 OY 566 SVKRLRRKRMNFVYCLMVSQGVPMFYMGDEYHTRKGNNTYCHDSYVYFPRMDK-EG 624  
 DB 541 SVKRLRRKRMNFVYCLMVSQGVPMFYMGDEYHTRKGNNTYCHDSYVYFPRMDK-EG 600  
 OY 625 YSELHRCFLMFKFKKCEGGLDEFTAKRQIMHSHOGKQDMDSNSRFVAFSMKDERO 684  
 DB 601 SSDLRFPCRLMTEFRKECSLGEDEFTAKRQIMHSHOGKQDMDSNSRFVAFSMKDERO 660  
 OY 685 GEIYVAFNTSHLPAYVELPERAGRMEPVYDGRKAPYFLTDPLDRALTIHOFSHFLY 744  
 DB 661 GEIYVAFNTSHLPAYVELPERAGRMEPVYDGRKAPYFLTDPLDRALTIHOFSHFLY 720  
 OY 745 SNLYPMLSYSYILVLRPDV 764  
 DB 721 SNLYPMLSYSYILVLRPDV 740

# RESULT 4

US-09-187-124-2  
 Sequence 2, Application US/09187124A  
 Patent No. 6255563  
 GENERAL INFORMATION:  
 APPLICANT: Emmertmann, Michael  
 APPLICANT: Kossmann, Jens  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES CODING FOR DEBRANCHING ENZYMES  
 TITLE OF INVENTION: FROM POTATO  
 FILE REFERENCE: GPEB8  
 CURRENT APPLICATION NUMBER: US/09/187,124A  
 CURRENT FILING DATE: 1998-11-05  
 EARLIER APPLICATION NUMBER: PCT/EP97/02292  
 EARLIER FILING DATE: 1997-05-06  
 EARLIER APPLICATION NUMBER: DE 196 18 125.9  
 EARLIER FILING DATE: 1996-05-06  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 2  
 LENGTH: 606  
 TYPE: PRT  
 ORGANISM: Solanum tuberosum  
 US-09-187-124-2

Query Match 62.0%; Score 2612; DB 4; Length 606;  
 Best Local Similarity 76.6%; Pred. No. 2;46-246;

Matches 458; Conservative 62; Mismatches 70; Indels 8; Gaps 3;

OY 171 ARG--NNCWPMAGMIPPLYSTFDEGDLPLRYPOKDLVIYEMHLRGFTKHDSSNVEHFG 228  
 DB 171 ARG--NNCWPMAGMIPPLYSTFDEGDLPLRYPOKDLVIYEMHLRGFTKHDSSNVEHFG 205

Db	3	ARGEDDCWCPMAGVNPASDOQFMEGDLKLPQRDLVYEMHVRGTFNNHSSSTKRG	62
QY	229	TFIGAVSKLDYLKELGVNCIELMPCHEHELEYSTSS-----KMFNGYSTINFESPT	283
Db	63	TYLGVEKLDHLKELGVNCIELMPCHEHELEYSYNSVLGDYKFNFGYSTVNFSPMG	122
QY	284	RYTSGGICNCRDAINERKTEVREHAKRGIEYILDVYNNHRAEGENENPILSFKVDVT	343
Db	123	RYSSGSLNCSGIGALINEFKYLYKREAHKRGIEYIMDVYNNHRAEGENENPILSFGRIDSV	182
QY	344	YVMLAPKGEFYNSGCGTNPNCNHHVYNOFIYDCLRYWVTEHHVGFPPDLASINTGSS	403
Db	183	FYTLAPKGEFYNSGCGTNPNCNHHVYNOFIYDCLRYWVTEHHVGFPPDLASINTGSS	242
QY	404	LMDDPVNVGALIEGEMITGTPLVYPPILDMISNDPILGGVYKLIAEADAGLYGVGFP	463
Db	243	SMNVNVYNSIDGMITGTPLVYPPILDMISNDPILGGVYKLIAEADAGLYGVGFP	302
QY	464	HMNVSEVNGKRYRDIVYRPFIKGTDFACGFACELCGSHLYOAGRKRWHSINFCADG	523
Db	303	HMGISEVNGKRYRDIVYRPFIKGTDFACGFACELCGSHLYOAGRKRWHSINFCADG	362
QY	524	FTLADLVYNNKYNLPNENNDGDNHNLNMGCEGEPARLSYRLKRRKMRNFVCLM	583
Db	363	FTLADLVYNNKYNLPNENNDGDNHNLNMGCEGEPARLSYRLKRRKMRNFVCLM	422
QY	584	VSQGVPMYMGDEYGTGKGNNTTCCHDSYVYFPMNDK-EQYSELHRCCLMTFRREC	642
Db	423	VSQGVPMYMGDEYGTGKGNNTTCCHDYINYPFMNDKDESSDPLFCGLMTFRREC	482
QY	643	EGLGLEDFPAKRLQWHGHPKPMSENSRFVAESMKDEROGELYVAFNTHLPAVEL	702
Db	483	ESLGDGFPFAERLQWHGHTPRTPMSETSRFVAFVTLVDKVLGELYAFNASHLPVTTL	542
QY	703	PERARRRHPVYDCKRAPYDLDLTPRALTIHQSFELXSNLYPMIYSSVITLV	760
Db	543	PERKPYRMQFPVDYCKRAPYDLDLTPPERETAAQYSHFLDANQYPMIYSSVITLV	600
US-09-346-237-9			
US-09-346-237-9			
Sequence 9, Application US/09346237A			
Patent No. 6265197			
GENERAL INFORMATION:			
APPLICANT: Bisgaard-Frantzen, Henrik			
APPLICANT: Svendsen, Allan			
TITLE OF INVENTION: Search Debbranching Enzymes			
FILE REFERENCE: 5629.200-US			
CURRENT APPLICATION NUMBER: US/09/346.237A			
CURRENT FILING DATE: 1999-07-01			
EARLIER APPLICATION NUMBER: PA 1998 00868			
EARLIER FILING DATE: 1998-07-02			
EARLIER APPLICATION NUMBER: 60/094.353			
EARLIER FILING DATE: 1998-07-28			
NUMBER OF SEQ ID NOS: 14			
SOFTWARE: FastSeq for Windows Version 3.0			
SEQ ID NO 9			
LENGTH: 713			
TYPE: PRT			
ORGANISM: Sulfolobus acidocaldarius			
FEATURE:			
NAME/KEY: PEPTIDE			
LOCATION: (1)..(713)			
OTHER INFORMATION: Isoamylase			
US-09-346-237-9			
Query Match			
29.2%; Score 1329.5; DB 4; Length 713;			
Best Local Similarity 38.1%; Pred. No. 3.9e-114;			
Matches 276; Conservative 101; Mismatches 238; Indels 109; Gaps			
21.			
60 GMPAPGATATL--AGGVNFAVYSGGATAALCLTFPEIDKADRYVEEVLPLDPLMNTGAV			
117			

```

Db      9  GEPPLPATWATWEEDDQVNFVLFSSNARKVELLTUYS-----QTRODEKKEITELRQRTGDL 63
QY      118  MHVFEIGELHMLLYGYRBDGTFAEPHCHGYLDVSNVVDPAKAAVISRGEYGVARGNCH 177
Db      64  MHVFEIPGLRPQLGYRYRGYGFKEEELRFENPKNKLIDYPAKAI-----NCLLLM 113
QY      178  PQMA-----CMIPLPYSTFPMWED---LPLRYPOKLVIY 209
Db      114  DDSVFGYKIGDONODLSFDERKDDKFIKPGVILNPF--FDMEDHEFFERRKIFPKOSIILY 171
QY      210  EMHLHGFKRRHDSNVNEH--PGTFIGAVS--KLDYKELGVNVCIELMPCHENELEYSTSS 266
Db      172  ETHIKGITKLNODLPENNRGFTFLASLADTMIDYKDGITTFVETLMPPOQVVDERFIYDG 231
QY      267  KMFNGYSTINFSPMTRYTYSGGIKNGCRDAINEKFTFVEBAHKRGIEVILDVVFNH7AE 326
Db      232  LKNWGYNPIVYFSPECKYSSSG---CLGNQVIEFKVLVNSLNHAGLEVIDVYVNH7AE 288
QY      327  GNENRPIISFEGVNTTYTHMLAPRGERY--NYSCCGTFPCNHPVYQFIVDCLRWY7TE 384
Db      289  GNHLCPPTLSFEGIDNSSYTHMLDPKPKRYIIDFTGTGNTLMLSHRVLQVLVDSLRYWLE 348
QY      385  MHVDFEFDLASIMTRGSSSLMDPVNVYVGAPRIEGDMITTGTPPLVPLIDMISNDPIIGV 444
Db      349  MHVDFEFDLASALAR-----QLYSV-----NMIST-----FFVAIQDDPIISQV 388
QY      445  KLIAEAMWDG--GLYOVGQFPMMYVSEWNGKRYDIBVROFIKGTGDFAGFRAECLGSPHL 503
Db      389  KLIAPPMVYGVGGYGVGNFPR--LMAENCKYRDTIRFRMGDPVPRYEELANRLGSPDL 446
QY      504  YVAGGRKRWHSINFCYCAHDGTTLADLVYTKKYYLPGENNRDGENHNLNSNGCEBEFA 565
Db      447  YAGSKTPEFASINYYTSHDGTTLDDLVSYNCKAHNAEARKLNNEGDMNENYSNNGCEBETN 506
QY      564  RLSYVRLEKRRQMRNFVCLVYNSGVPMFPMYDEGHTKNGGNNTYCHDSYVNTFRMDKE 623
Db      507  DSNILYCEKORNRNVITLVEFVSQIGIPMLGSDLEGKQKGNNAFCCODNETSWTDNMLDE 566
QY      624  QYSELHRCCLMTKTRKE-----CEGLGEDPFLAKRLQW---HGHPGKPRDSEN 671
Db      567  NNVRRHDEVRRLTNYYKAHPIFRARARYGCKLLGSP--LQDVTLKPKDQNEVVDSDVYKSP 625
QY      672  SRFVAFSMKDEROGEI-----YVAFNTHSLPAAVELPERAGRMHEPVDTGKP 719
Db      626  TWHIYIILEGSAIDBINYNGERIADDFLILNGASTNLKIKVPH--GKWEVLV--HP 679
QY      720  APYD 723
Db      680  YPHE 683

RESULT      6
US-09-346-237-10
? Sequence 10, Application US/09346237A
? Patent No. 6265197
? GENERAL INFORMATION:
? APPLICANT: Bisgaard-Frantzen, Henrik
? APPLICANT: Bisgaard, Allan
? TITLE OF INVENTION: Starch Debranching Enzymes
? FILE REFERENCE: 5629.200-US
? CURRENT APPLICATION NUMBER: US/09/346.237A
? CURRENT FILING DATE: 1999-07-01
? EARLIER APPLICATION NUMBER: PA 1998 00868
? EARLIER FILING DATE: 1998-07-02
? EARLIER APPLICATION NUMBER: 60/094,353
? EARLIER FILING DATE: 1998-07-28
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 10
? LENGTH: 718
? TYPE: PRT
? ORGANISM: Sulfolobus sulfataricus
? FEATURE:

```

```

QY 57 VLAAMPRLGATLALAGVNFVAVYSGGATTAALCLFT--PEDLKKADRVTEEVLPDLIMNRTG 115
Db 14 VMWRPRLPLGATMDGLGNVNAISQHNAAVELVLFHPDDPADSRTE-----VTERTG 67
QY 116 NMVHVFIEGELHNNMLYGYREDGTTPAPCHGYLDVSNVVDPAKAVISRGEYGPARGNN 175
Db 68 PIMHVYLPGLRPGQLGYRVRYGYPRIPEGRHFNPNKKVLDLPYAKAI-----GNPLK--- 118
QY 176 CRPM-----AGHIFPL-----PYS-----TFMEGDLPLRTQOKLDVITYEMH 212
Db 119 -MHDSLEGYKIDPAGDLSFSEEDSAPYAPLGAIVVEGCFEMGDDRPRIEMEDIIYETH 177
QY 213 LRGTCKDSSNVHP-----GTFIAVSK--LDYCKELGVNICTELMPCHEFNLELST 265
Db 178 VKGJTK-----LHPVPEPLRGTYLDTLCEPVLHNLKQLGYTTTQLLPHAKYHNRHLV 233
QY 264 SSSKMNFWGYSTINFESPMTRY--TSGIKKCGRDAINEFTFYEAHKRGIEVLDVFN 322
Db 232 ERLGRNMGVNPCLCYFAPEPEYATNGPI-----SAVREPKMVRALHAAGEVIVDVYN 286
QY 323 HTAAGNENPLSKGVNDNTTYMLAKGEFY--NYSGCCGTFPCNIPVYRQFLVDCRLY 386
Db 287 HTGGGVLGPLTSPRGIDNRYYKADNPNRFLVDYGTGNTLDGVPNYIQLIMDSLRY 346
QY 381 WTEMHNVDFEFEDLASINTRGSSLMDEVNYYGAPIEGDMITTTGPLVTPPLIMDISNDPI 440
Db 347 WTEMHNVDFEFEDLAALAR--ELIYD-----DMIST-----PFYVIQDDPV 386
QY 441 LGGVKLIAEAWDAG--GLYOVGQPRHNMVWSEBNGKYRDIYRQFIKGTDFAGGFAECLCG 495
Db 387 LSGYKLLAEPMDVPGRYGYQHVP--WQ--WTEMNKRYADAVRRFRGDRGLNGEATFAG 444
QY 500 SPHLIYAGGKRPWHSINFCVCHDGFTLADLVTKKTYNLPRGENNRDGENHNLSWNGCEE 555
Db 445 SSDLYERSGRPFASINFTVAHDFLTLEDLYSTYKHNENALEGNRQGDENYSTNGCVE 504
QY 560 GEFARLSYKRLRKQKQMNFFCLMVSGGVPRFYMGDEYGTGKGNNTTYCHDSVYVTFW 619
Db 505 GPTODPVSALCREALKRSLSTLTPLSGVPVPLLGGLDLSRTQHNHNAAYCQDNEISYWN 564
QY 620 D---KKEDYSLEHAFCCLMTKFKRECEGLGLEDPYAKRLQMHCHOP-----GKP-- 666
Db 565 QLDTRKQOFLF---FVQOTIMFKRQHSFRNRHPLTLP---NGRRRBSLYAPREGPRM 618
QY 667 --DWSNSRFAVESM-----KDE-----ROGEIYVAFNTSHLPVAVELPERAG-- 707
Db 619 HEDWT--NPELTAFLGLLHGALIOGTDEHGRRPFDRTFLILENNSEAVPVVVEVCSGK 677
QY 708 -RRREPV-----VDTRKPARPYDLETLDDLRALTI 736
Db 678 PHHVEVVPVFORNVEPPTCAPGETLS--LRPGVLTV 711

```





Query Match 19.2% Score 810; DB 1; Length 777;  
Best Local Similarity 32.6%; Pred. No. 4.4e-70;  
Matches 257; Conservative 96; Mismatches 271; Indels 164; Gaps 38;

```
QY 42 EPVAEDRYALGACRVL-----AGMPA-----PLGATALAGVN--FAVYSGATA 85
DB 2 DPHAPQROSGORLALALALACALSPAHAAIDAQOLGARDAQAQANLAFRYVSSRATR 61
QY 86 AALCLF--TPED-----LKADRYTE---EVLPLDMNRTGNVHWVFTEGELHNNLYG 132
DB 62 VEFLYKNPTGSOEVARLALSKDPATQVMSLSLPTSTIKNTYG-----ITGAVY---YG 112
QY 133 YRFDGTFAPH-----CGHYLDVSN-----VVDPRYAK-----AVISRGE 166
DB 113 YRAMGPMMPYDAWTKSATGFSVDVNDAGNREPNKLLDPYAREISQDPNTATACAGT 172
QY 167 -YGVPARGN-----CWPOMAGMIPLPYSTFDEMGDLPLRYPOKDLVYEMHLRGFTKHD 220
DB 173 IYATGAHRNKDGLCASKGIALAADATSV---GSKPTR-ALKDEVITYEVHVGRLTRND 227
QY 221 SS-NVEHPOTFIGAVSKLDYKELGVNCEIEMPCHEF---NELEYSTSSKMNEMGYST 275
DB 228 DSVPAERGTYYKGAARKAAALALGVTAVEFLPVQETQNDQNDVD--PNSTAGDNWGYMT 286
QY 276 INFESPMTRY---TSGGKNGCRDAINEFKTFVREAHKRGIEVLIDVVENHTAEGNENG 331
DB 287 LNYFAPDRRYAYDKSAG-----PTREWKAMVAFHDAGIKVYIDVYVYNTGEG---G 336
QY 332 P-----ILSFKGVNDTTYMLAPKGEF--YNSGCGTFCNHPVROFIYDCLR 379
DB 337 PWSGTDLGVNLLSFRLDNPAYSLSSDYKYPMDNTGVGNNYTRHPIDQNLIVDSLA 396
QY 380 YWVTEMHVDFRFDLASIMTRGSSLMDPVNVYGAPIEGDMITTTGTPPLVPLDIMISNDP 439
DB 397 YWRALGVDFRFDLASYL--GNSCOHCCFNPKDKNDSGNALNR-----IYAEILPPRP 446
QY 440 ILG--GVKLIAPMADAG--LYOVGOFPHMNWSEWNGKRYDIR--QFIKTDGFA--GGF 493
DB 447 AAGGAGADILAPWALIGNSYOVGGFPAG--WAEWNGLYRDLARRKKNKLGEVTVTPGL 504
QY 494 AECLGSPHLIYQAGGRKRWHSINFCADHDFTLADLVYNNKKYV---LPNGENNRDGENH 550
DB 505 ATRFAGSNDLYGDDGRKRWHSINFVVAHDGFTLNDLYAYNDQNNQPMWYGPSD--GGEDEH 563
QY 551 NLSWNGEGEGEFARLSVRLRKROMNFEVCLMVSGVPMFYMGEYGHTKGNNNTYCH 610
DB 564 NLSWNG-----GIVAEOKRAKARTGLALMLLSAGVPMITGGDBALRTQCGNNNTYUL 615
QY 611 DSYVNTFRMDKROYSELRFCCLMTKFRKECEGLLEDFPTA-----KRLQMHGHO 662
DB 616 DSAANMLVWSRSALAEADHETTKRLIAFRKAHPALRPANFYASDPTNGVMQLRWF--- 672
QY 663 PGKPDWSE-----NSRFVAFSMKDEROGE-----IYVAFNTSHLPAVELLP--ERAGR 708
DB 673 --KPDGADASAYFNAGDNHMLAWRIDSEFGDSASAIYVAVNGSAGVDFKLPWPGTGK 730
QY 709 RMEPVYDT 716
DB 731 QWRYVTDIT 738
```

RESULT 12  
PCT-US95-09323-2  
Sequence 2, Application PC/TUS9509323

GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Novel Isoamylase Gene, Compositions  
TITLE OF INVENTION: Containing It and Methods of Using Isoamylases  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/09323  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/281902  
FILING DATE: 28-JUL-1994  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 777 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

PCT-US95-09323-2

Query Match 19.2% Score 810; DB 5; Length 777;  
Best Local Similarity 32.6%; Pred. No. 4.4e-70;  
Matches 257; Conservative 96; Mismatches 271; Indels 164; Gaps 38;

```
QY 42 EPVAEDRYALGACRVL-----AGMPA-----PLGATALAGVN--FAVYSGATA 85
DB 2 DPHAPQROSGORLALALALACALSPAHAAIDAQOLGARDAQAQANLAFRYVSSRATR 61
QY 86 AALCLF--TPED-----LKADRYTE---EVLPLDMNRTGNVHWVFTEGELHNNLYG 132
DB 62 VEFLYKNPTGSOEVARLALSKDPATQVMSLSLPTSTIKNTYG-----ITGAVY---YG 112
QY 133 YRFDGTFAPH-----CGHYLDVSN-----VVDPRYAK-----AVISRGE 166
DB 113 YRAMGPMMPYDAWTKSATGFSVDVNDAGNREPNKLLDPYAREISQDPNTATACAGT 172
QY 167 -YGVPARGN-----CWPOMAGMIPLPYSTFDEMGDLPLRYPOKDLVYEMHLRGFTKHD 220
DB 173 IYATGAHRNKDGLCASKGIALAADATSV---GSKPTR-ALKDEVITYEVHVGRLTRND 227
QY 221 SS-NVEHPOTFIGAVSKLDYKELGVNCEIEMPCHEF---NELEYSTSSKMNEMGYST 275
DB 228 DSVPAERGTYYKGAARKAAALALGVTAVEFLPVQETQNDQNDVD--PNSTAGDNWGYMT 286
QY 276 INFESPMTRY---TSGGKNGCRDAINEFKTFVREAHKRGIEVLIDVVENHTAEGNENG 331
DB 287 LNYFAPDRRYAYDKSAG-----PTREWKAMVAFHDAGIKVYIDVYVYNTGEG---G 336
QY 332 P-----ILSFKGVNDTTYMLAPKGEF--YNSGCGTFCNHPVROFIYDCLR 379
DB 337 PWSGTDLGVNLLSFRLDNPAYSLSSDYKYPMDNTGVGNNYTRHPIDQNLIVDSLA 396
QY 380 YWVTEMHVDFRFDLASIMTRGSSLMDPVNVYGAPIEGDMITTTGTPPLVPLDIMISNDP 439
DB 397 YWRALGVDFRFDLASYL--GNSCOHCCFNPKDKNDSGNALNR-----IYAEILPPRP 446
QY 440 ILG--GVKLIAPMADAG--LYOVGOFPHMNWSEWNGKRYDIR--QFIKTDGFA--GGF 493
DB 447 AAGGAGADILAPWALIGNSYOVGGFPAG--WAEWNGLYRDLARRKKNKLGEVTVTPGL 504
QY 494 AECLGSPHLIYQAGGRKRWHSINFCADHDFTLADLVYNNKKYV---LPNGENNRDGENH 550
DB 505 ATRFAGSNDLYGDDGRKRWHSINFVVAHDGFTLNDLYAYNDQNNQPMWYGPSD--GGEDEH 563
QY 551 NLSWNGEGEGEFARLSVRLRKROMNFEVCLMVSGVPMFYMGEYGHTKGNNNTYCH 610
DB 564 NLSWNG-----GIVAEOKRAKARTGLALMLLSAGVPMITGGDBALRTQCGNNNTYUL 615
QY 611 DSYVNTFRMDKROYSELRFCCLMTKFRKECEGLLEDFPTA-----KRLQMHGHO 662
DB 616 DSAANMLVWSRSALAEADHETTKRLIAFRKAHPALRPANFYASDPTNGVMQLRWF--- 672
QY 663 PGKPDWSE-----NSRFVAFSMKDEROGE-----IYVAFNTSHLPAVELLP--ERAGR 708
```

Db 673 --KPDGAQADSAVFNAGADNHALAWRIDGSEFGDSASAIYVANGMSGAVDFKLPWCTGK 730  
Qy 709 RMEPVVD 716  
Db 731 QWYRVDT 738

## RESULT 13

US-08-476-519-11  
; Sequence 11, Application US/08476519  
; Patent No. 5750876  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Krohn, Bradley M.  
; TITLE OF INVENTION: No. 5750876e1 Isoamylase Gene, Compositions  
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Grace L. Bonner, Monsanto Company, Bldg  
; STREET: 700 Chesterfield Parkway No. 5750876ch  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,519  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/281902  
; FILING DATE: 28-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bonner, Grace L.  
; REGISTRATION NUMBER: 32,963  
; REFERENCE/DOCKET NUMBER: 38-21(13577)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-7286  
; TELEFAX: (314)537-6047  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 746 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-476-519-11

Query Match 19.2%; Score 807; DB 1; Length 746;

Best Local Similarity 33.0%; Pred. No. 8.1e-70;  
Matches 248; Conservative 95; Mismatches 258; Indels 150; Gaps 36;

Qy 65 LGATLALAGVN--FAVYSGATAALCLF--TPED-----LKADRYTE---EVLDP 109  
Db 8 LGARDAQAANLAFRYSSRATREVEFLKKNPFGSOEVARLALSKPATOWMSLSLPTST 67  
Qy 110 LMRGTGNWVHPIEGELHMLGYREDGTFAPI-----CGHYLDVSN----- 151  
Db 68 IKNTYG-----ITGAVY---YGYRAMGNWMPYDAWTKGATGFSVDNAGNRPNNK 118  
Qy 152 VVVDPIAK-----AVISRGE--IGVPARGN-----CWPQAGATPLPSTFDWEGDL 197  
Db 119 LLIDYAREISODPNTATCADGTIYATGAHRRKDSGLCASKIALAADATSV---GSK 174  
Qy 198 PLRYOKDLVIYEMHLRGFTKHDSS--NVEHPTGFTIGAVSKLDVLEKLVNCFILMPCHER 256  
Db 175 PIR-ALKDEVITYEVHVRGLTRNDSDVPAERGTGYKGAARKAALALAGTAAVEFLVQET 233

Qy 257 -----NELEYSTSSSKMFMGTYSTINFSPMTRY-----TSGGIRKNGRDAINEKTFVREA 308  
Db 234 QNDQNDVD--PNSYAGDNYWMTNLNPAAPRRYAYDSAG-----PTREWMYKAF 285  
Qy 309 HKRGIEVLDVFNHPTAEKNENGP-----ILSPKGVNDTTYMLAPKGEF--YNY 356  
Db 286 HDAGIKYIDVYNNHTEG---GPMSTGDLSTYNLSTFRLDNPATYSSSDYKTFWDN 342  
Qy 357 SGGCNFTNCNHPVYRQPIVDCLRYWTEMHVDEGRFPLASIMRGSILMPVNYGAPIE 416  
Db 343 TGVGNVNTNHPILAQNLIYDSLAWRDALGVDFRPLASVL--GNSCQHCFCFEDNDS 400  
Qy 417 GDMITGTPLVTPPLIDMISNDPLG--GVKLLAEANDAG--LYQVQCPHMYWSENG 473  
Db 401 GNALNR-----YVALPFRPAGAGADLLAEPAIAGNSYQVGFAPG--WAEWNG 450  
Qy 474 KYBDIVR--OFIKGTQDFA--GGEFECGSPHLXQACGRKPMWHSINFCVCHDGTLDLY 530  
Db 451 LTRDALRKKONKLGVEVTPGTLTRPAGSNDLGDGGRKPMWHSINFCVHADGFTLADLY 510  
Qy 531 TYNKKYV---LPNGENNRDENHNLWNCGEGRFARLSYRLKRRMRNFFVCLAVYSG 587  
Db 511 AYNDKQNNQMPYGPSPD--GGEEDHLSWNG-----GIVAEQRKAKRGLALMLISAG 561  
Qy 588 VPMFTMGDEYGHYKGGNNNTYCHDSYVNYFRMDKKEQYSELARFCCLMTFRKECEGL 647  
Db 562 VPMITGDEALRTOFGNNNTYNLDSANWLYWSRSLAEADHETVTKRLIAFRKAPALRP 621  
Qy 648 EDPEPTA-----KRLQWGHQGRKPMSE-----NSRVAVSMKDERGE----- 686  
Db 622 ANFYSDTNGVNEQLRMP-----KPDGAQADSAVFNAGADNHALAWRIDGSEFGDSASA 676  
Qy 687 IYVAFNTSHLPAVELP--ERAGRMEPVVD 716  
Db 677 IYVAVNMSGAVDFKLPWCTGKQWYRVDT 707

## RESULT 14

PCT-US95-09323-11  
; Sequence 11, Application PC/TUS9509323  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions  
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases  
; NUMBER OF SEQUENCES: 11  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/09323  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/281902  
; FILING DATE: 28-JUL-1994  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 746 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-09323-11

Query Match 19.2%; Score 807; DB 5; Length 746;

Best Local Similarity 33.0%; Pred. No. 8.1e-70;  
Matches 248; Conservative 95; Mismatches 258; Indels 150; Gaps 36;

Qy 65 LGATLALAGVN--FAVYSGATAALCLF--TPED-----LKADRYTE---EVLDP 109  
Db 8 LGARDAQAANLAFRYSSRATREVEFLKKNPFGSOEVARLALSKPATOWMSLSLPTST 67







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:05:30 ; Search time 44.53 seconds

(without alignments)  
1648.602 Million cell updates/sec

Title: US-09-674-817A-2

Perfect score: 4212

Sequence: 1 SGPAFLRLRRMPPNATACKGV.....SNLYPMLSYSSVILVLRDPV 764

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3568.5	84.7	789	2	T01321
2	3527.5	83.7	818	2	T02231
3	2895.5	68.7	783	2	B84823
4	1638.5	38.9	707	2	S77094
5	1538.5	36.5	702	2	B85091
6	1388	33.0	746	2	S74546
7	1305.5	31.0	664	2	B86539
8	1305.5	31.0	664	2	B72084
9	1300	30.9	666	2	E71565
10	1268	30.1	720	2	B75540
11	1264	30.0	666	2	B81717
12	1225	29.1	721	2	A70764
13	1211	28.8	718	2	S73088
14	1149.5	27.3	716	2	C83375
15	1123.5	26.7	698	2	C96001
16	1101.5	26.2	603	2	AC0674
17	1014	24.1	659	2	A64119
18	1005.5	23.9	657	2	UC7767
19	1002	23.8	657	2	D91163
20	1001	23.8	657	1	BVECGX
21	997	23.7	657	2	B86009
22	980.5	23.3	654	2	AG0495
23	967.5	23.0	662	2	AG0495
24	927.5	22.0	656	2	D82388
25	890	21.1	882	2	F86164
26	815.5	19.4	651	2	B98229
27	815.5	19.4	651	2	AD3057
28	803.5	19.1	776	2	A37035
29	798.5	19.0	776	2	S13470

30	527.5	12.5	718	2	G69585	pullulanase amyX -
31	523	12.4	843	2	H72204	pullulanase - Ther
32	519.5	12.3	720	2	G97229	pullulanase (Impor
33	505.5	12.0	717	2	H84057	pullulanase amyX l
34	420	10.0	759	2	A98000	alpha-dextrin endo
35	418	9.9	759	2	D95129	pullulanase, proba
36	408	9.7	1072	2	A84112	alkaline amylopull
37	395.5	9.4	1256	2	G97902	alpha-amylose (EC
38	394.5	9.4	1280	2	E95031	alkaline amylopull
39	370.5	8.8	561	2	S73087	alpha-amylose (EC
40	354	8.4	691	2	D90592	hypothetical prote
41	344	8.2	964	2	S51324	pullulanase - spin
42	335	8.0	558	2	JC5135	alpha-amylose (EC
43	327.5	7.8	910	2	D75524	alpha-dextran endo
44	327	7.8	966	2	T03760	alpha-dextrin endo
45	325.5	7.7	904	2	T04377	probable alpha-dex

#### ALIGNMENTS

##### RESULT 1

probable isoamylase (EC 3.2.1.68) su1 - maize

C:Species: Zea mays (maize)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 29-Oct-1999

C:Accession: T01321

R:Park, S.Y.; Kim, Y.H.

Plant Physiol. 115, 1731, 1997

A:Title: Cloning and nucleotide sequencing of a cDNA encoding flavanone 3-hydroxylase

A:Reference number: Z14292

A:Accession: T01321

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-789 <PAR>

A:Cross-references: EMBL:AF030882; NID:q2623247; PIDN:AA897167.1; PID:q2623248

C:Genetics:

A:Gene: sugary1

A:Insertions: 121/3; 183/3; 214/3; 265/3; 291/2; 338/3; 374/3; 403/2; 427/2; 469/3; 491

C:Function:

A:Description: starch debranching enzyme

C:Superfamily: glyx protein

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match

Best Local Similarity 84.7%; Score 3568.5; DB 2; Length 789;

Matches 654; Conservative 44; Mismatches 65; Indels 19; Gaps 6;

OY	1	SGPAFLRLRRMPPNATACKGV-----CAAYEAATKVEDEGEDEBPV	44
DB	9	SSPRLDLPVAGRWRAVGRHPN-VAGIGRGRLSLHAAAPRAVAQAEEDDDDDDEV	67
OY	45	AEDRYAAGACRVLAGMPAPRGATAGLVGFAYTSSGATPAACTPREDLAKDRYTEE	104
DB	68	AEEFALGACRVLAGMPAPRGATAGLVGFAYTSSGATPAACTPREDLAKDRYTEE	127
OY	105	VPLDPLNRTGNNVHVFIEG-ELHNLMLXGYFEDGTFAPHCGLYDVSNNVVDPAKAYIS	163
DB	128	VPLDPLNRTGNNVHVFIEG-ELHNLMLXGYFEDGTFAPHCGLYDVSNNVVDPAKAYIS	187
OY	164	RGEYGVAPRGSCWPAQAGMIPRYSTFDEMDGLPLRYPOKDLVYEMHLNGFTKHDSSN	223
DB	188	RGEYGVAPRGSCWPAQAGMIPRYSTFDEMDGLPLRYPOKDLVYEMHLNGFTKHDSSN	247
OY	224	VENHGTFTGAVSKDYLKEGLVNCIELMPCHFEFNELEYSTSSKAMNTGISTYTFESPMT	283
DB	248	TKHGTFTGAVSKDYLKEGLVNCIELMPCHFEFNELEYSTSSKAMNTGISTYTFESPMT	307
OY	284	RYTSGGIRKCGDAINEFKTFVREAHKRGIVLIDVYFNHTAENNGPILSFGVDVDTT	343
DB	308	RYTSGGIRKCGDAINEFKTFVREAHKRGIVLIDVYFNHTAENNGPILSFGVDVDTT	367

QY 344 YVLAARKGEFYNSGCCNTFNCNHPVYVROFIVDCRLRYVTEMHVDSGRFDLASTINTRGSS 403  
 DB 368 YVLAARKGEFYNSGCCNTFNCNHPVYVREIVDCLRYVTEMHVDSGRFDLASTINTRGSS 427  
 QY 404 LMDPVNVYGPRIEGDMITTTGTPVLPPLIMISNDPILGGVKTIAEAMDAGGLYQVGOFP 463  
 DB 428 LMDPVNVYGPRIEGDMITTTGTPVLPPLIMISNDPILGGVKTIAEAMDAGGLYQVGOFP 487  
 QY 464 HMMVSEMNKGYRDIYVROFVKTGDFAGFAECICGSPHLYOAGGRKPMHSINFCVCAHDG 523  
 DB 488 HMMVSEMNKGYRDIYVROFVKTGDFAGFAECICGSPHLYOAGGRKPMHSINFCVCAHDG 547  
 QY 524 FTLDLVYTKKYNLPNGENNRDGENHNLNMGCEBGEFARLSYRLKRRKMRNPFVCLM 583  
 DB 548 FTLDLVYTKKYNLPNGENNRDGENHNLNMGCEBGEFARLSYRLKRRKMRNPFVCLM 607  
 QY 584 VSGVPMFYMGDEYHTGKGGNNNTYCHDSYVNYFRMDK -EQYSELHRCCLMTKFRKEC 642  
 DB 608 VSGVPMFYMGDEYHTGKGGNNNTYCHDSYVNYFRMDK -EQYSELHRCCLMTKFRKEC 667  
 QY 643 EGIAGLEDEPTAKRLQMHQHPKPMSENSRFAVAFSMKDERGGEIYVAFTNSHLPAVVEL 702  
 DB 668 ESIAGLEDEPTAKRLQMHQHPKPMSENSRFAVAFSMKDERGGEIYVAFTNSHLPAVVEL 727  
 QY 703 PERAGRRMEPVYDTGKRAPYDFLTDDLPRALTIHQFSHFLYSNLXPMLSYSSVILVLRP 762  
 DB 728 PERAGRRMEPVYDTGKRAPYDFLTDDLPRALTIHQFSHFLYSNLXPMLSYSSVILVLRP 787  
 QY 763 DV 764  
 DB 788 DV 789

RESULT 2  
 T02231  
 Probable isoamylase (EC 3.2.1.68) su1 - maize (fragment)  
 C:Species: Zea mays (maize)  
 C>Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 29-Oct-1999  
 C:Accession: T02231  
 R:James, M.G.; Robertson, D.S.; Myers, A.M.  
 Plant Cell 7, 417-429, 1995  
 A:Title: Characterization of the maize gene sugary1, a determinant of starch composition  
 A:Reference number: 214630; MUID:95290995  
 A:Accession: T02231  
 A>Status: preliminary: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-818 <JMAN>  
 A:Cross-references: EMBL:U18908; NID:g758725; PIDN:AAA91298.1; PID:g758726  
 A:Experimental source: endosperm  
 C:Genetics:  
 A:Gene: sugary1  
 A:Map position: 45  
 C:Function:  
 A:Description: starch debranching enzyme  
 A:Superfamily: glyx protein  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 83.7%; Score 3527.5; DB 2; Length 818;  
 Best Local Similarity 83.0%; Pred. No. 2e-263; Indels 19; Gaps 6;  
 Matches 649; Conservative 44; Mismatches 70;  
 QY 1 SGPAPRLR-----RM-----RPNATGKGVGEV-----CAAVVEAATVVEDEGEDEPY 44  
 DB 38 SSPRLAVPACGRWRACVGRNR-VAGLGRGRSLHAAARPAVAEVAQAEEDDDDEY 96  
 QY 45 AEDRYALGACRVLGMPAPLGCATLALAGVNFVAVSGATAALCLFTPEDEKADRVTEE 104  
 DB 97 AEFRLGACACVLAGMPAPLGCATLALAGVNFVAVSSGASASLSIFAGDGLKADRVTEE 156  
 QY 105 VPLDPLMNTGNVWVHFIIG-ELHNNLXYGRPDGTFAPHCGRHLDVSNVVDYAKAVIS 163  
 DB 157 VPLDPLMNTGNVWVHFIIGDELHGMLCGRDGVFAPERGQYDYVSNVVDYAKAVIS 216

QY 164 RGEYVAPARGNMCWPMAGMTPLPYSTFDEMGDPLRYRPOCDLVYEAMHKGFTKHDSSN 223  
 DB 217 RGEYVAPARGNMCWPMAGMTPLPYSTFDEMGDPLRYRPOCDLVYEAMHKGFTKHDSSN 276  
 QY 224 VEHPTFIGAASKIDYKELKELVNCIELMPCHEFNELLEYSTSSSKNFWGYSTINFSPMT 283  
 DB 277 TKHGTIYGANASKIDHKLKELVNCIELMPCHEFNELLEYSTSSSKNFWGYSTINFSPMA 336  
 QY 284 RYTSGGIKNCGRDAINEKTVFREAHKRGIVYILDVFNHTAEGNENPILSFKVDVMT 343  
 DB 337 RYSSGGIRDSGCAINERKAFVREAHKRGIVYILDVFNHTAEGNENPILSFRCIDMST 396  
 QY 344 YVLAARKGEFYNSGCCNTFNCNHPVYVROFIVDCRLRYVTEMHVDSGRFDLASTINTRGSS 403  
 DB 397 YVLAARKGEFYNSGCCNTFNCNHPVYVREIVDCLRYVTEMHVDSGRFDLASTINTRGSS 456  
 QY 404 LMDPVNVYGPRIEGDMITTTGTPVLPPLIMISNDPILGGVKTIAEAMDAGGLYQVGOFP 463  
 DB 457 LMDPVNVYGPRIEGDMITTTGTPVLPPLIMISNDPILGGVKTIAEAMDAGGLYQVGOFP 516  
 QY 464 HMMVSEMNKGYRDIYVROFVKTGDFAGFAECICGSPHLYOAGGRKPMHSINFCVCAHDG 523  
 DB 517 HMMVSEMNKGYRDIYVROFVKTGDFAGFAECICGSPHLYOAGGRKPMHSINFCVCAHDG 576  
 QY 524 FTLDLVYTKKYNLPNGENNRDGENHNLNMGCEBGEFARLSYRLKRRKMRNPFVCLM 583  
 DB 577 FTLDLVYTKKYNLPNGENNRDGENHNLNMGCEBGEFARLSYRLKRRKMRNPFVCLM 636  
 QY 584 VSGVPMFYMGDEYHTGKGGNNNTYCHDSYVNYFRMDK -EQYSELHRCCLMTKFRKEC 642  
 DB 637 VSGVPMFYMGDEYHTGKGGNNNTYCHDSYVNYFRMDK -EQYSELHRCCLMTKFRKEC 696  
 QY 643 EGIAGLEDEPTAKRLQMHQHPKPMSENSRFAVAFSMKDERGGEIYVAFTNSHLPAVVEL 702  
 DB 697 ESIAGLEDEPTAKRLQMHQHPKPMSENSRFAVAFSMKDERGGEIYVAFTNSHLPAVVEL 756  
 QY 703 PERAGRRMEPVYDTGKRAPYDFLTDDLPRALTIHQFSHFLYSNLXPMLSYSSVILVLRP 762  
 DB 757 PERAGRRMEPVYDTGKRAPYDFLTDDLPRALTIHQFSHFLYSNLXPMLSYSSVILVLRP 816  
 QY 763 DV 764  
 DB 817 DV 818

RESULT 3  
 B84823  
 Probable isoamylase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: B84823  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: B84823  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-783 <STO>  
 A:Cross-references: GB:A8002093; NID:g2088646; PIDN:AA895278.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g39930  
 A:Map position: 2  
 C:Superfamily: glyx protein

Query Match 68.7%; Score 2895.5; DB 2; Length 783;  
 Best Local Similarity 69.3%; Pred. No. 9.2e-215;  
 Matches 516; Conservative 84; Mismatches 128; Indels 17; Gaps 5;

```

Oy 28 VEATKRVDEGE---DEPVAEDRYALGACRVLGAMPALGATAGAGVNFVSSGA 83
Db 42 ISAKRRSNEAENIAYVEEPLKSDREFISD-----GLSPSPGPTPRDDGVNFVSTYNS 95
Oy 84 TAAALCLFTEPDJLKADRYTEEYPLDPLMNRGTGNVHVPIEGELHNMLYGREDGTAPAPIC 143
Db 96 VSATCICLSLSDLRQNKVTEEIQDPSRNRTGVMHVFLRGDKOKMLYGRFGKRSPEE 155
Oy 144 GHYLDVSNVVDPIYAKAVISRGEGYVARGNCCWPMACMIPLPYSTFMEGDLPLRYPO 203
Db 156 GHYDSSNLLDPIYAKAISRDEFGVIGPDNCPQMACMVPTRESEFMEGDMHLKLPQ 215
Oy 204 KDLVYTEMHRLKDHSSNVHEPGFICGAVSKLDYKELGVGICLMPCHENLEYSY 263
Db 216 KDLVYTEMHVRGFTRESKIEPPTGYQVAEKLDHLKELGICLMPCHENLEYSY 275
Oy 264 SSS-----KMFVGYSTINEFSPMRYTSGIKNCGRDAINEKTEFVREAHKRGIEVILD 318
Db 276 YNTILGDHRVNFVWGYSITFEFSPMIRYASASSNPFAGRAINEKILVKEAHKRGIEVIMD 335
Oy 319 VYFNHTAEGNENGPILSEFGVNDITYYMLAPKGEFYNYSGCGTENCNHPVYRQFTYDCL 378
Db 336 VYLNHTAEGNEKGPITSEFGVNSVYVYMLAPKGEFYNYSGCGTENCNHPVYRQFTYDCL 395
Oy 379 RYVWTEMHVDFRFDLASTINTRGSSSLMDPVNYGARIIEGDMITGTPLTPIIDMISND 438
Db 396 RYVWTEMHVDFRFDLGSITMSRSSSLMDANVYGADEVEDLTTGTPISCPVIDMISND 455
Oy 439 PIIIGGVKLLAEAMDAGGLYQVGOFPMMNVMSENGKYRDIIVROFIGTDFAGGAFACLC 498
Db 456 PILRQVKKLLAEAMDAGGLYQVGMFPMHGMISENKGKFRDVAKRFICTGDFSAFAECCLC 515
Oy 499 GSPHLVYQAGRRPWHISINFCVCAHDEFTLADLYTNKKYNLPNGENNRDGENHNLISWNGE 558
Db 516 GSPNLVYQ--GGRPMWHSINFCVCAHDEFTLADLYTNKKYNLANGEENNRDGENHNYSNWNGE 574
Oy 559 EGEEFARLSYKRLRKROMRNFYCLAMYSQGVPMFYNGDEGHTKGGNNNTYCHDSYVNYR 618
Db 575 EEDFASISYKRLRKROMRNFYCLAMYSQGVPMFYNGDEGHTKGGNNNTYCHDSYVNYR 634
Oy 619 WDKKQO--YSELHRCCLMTRFKRECEGLDEDEPTAKRLQMHGHOGRKPDMSSENSRFVAF 677
Db 635 WDKKEAHSDFRFRCLILKFRDECESLINDPTAKRLQMHGHLAEIPWMSSTSEFVAF 694
Oy 678 SMKDERGCEIYVAFNTSHLPAVELEPERAGRWEPVDTGKPARVYELTDLDRALTIH 737
Db 695 SLVDSVKKRIEYVAFNTSHLATVSLPNRPGYRWEPVDTGKPARVYELTDLDRALTIH 754
Oy 738 QESHFLYSNLYPLMSYSSIVLVRP 762
Db 755 QYRHLFADANVYPLMSYSSITILLSP 779

```

## RESULT

4

glycogen operon protein (EC 3.2.1.-) glgX-2 - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein slr1857

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S77094

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

5

A:Reference number: S74322; MUID:97061201

A:Accession: S77094

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-707 <KAN>

A:Cross-references: EMBL:D90908; GB:AB001339; NID:g1652725; PIDN:BAAL7652.1; PID:g165273

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:  
A:Gene: glgX-2  
C:Superfamily: glgX protein  
C:Keywords: glycosylase; hydrolase

Query Match 38.9%; Score 1638.5; DB 2; Length 707;

Best Local Similarity 49.1%; Pred. No. 5.6e-118;

Matches 342; Conservative 77; Mismatches 216; Indels 61; Gaps 16;

```

Oy 56 RVLGAPAPLPGATAGAGVNFVSSGATAAALCFTEPDJLKADRYTEEYPLDPLMNRGTG 115
Db 16 KILGCGPFRPGATAGAGVNFVSSGATAAALCFTEPDJLKADRYTEEYPLDPLMNRGTG 115
Oy 116 NVV--HVFIEGELHNMLYGREDGTAPAPIC 174
Db 72 NVYCMVVFDDFENLEEGYMEGRPNPQOGHMPDPSVLLDPAKAVYSGHDVGTQPNMD 131
Oy 175 NCMPPMAGMPLPYSTFMEGDLPLRYPOKDLVYEMHRLKDHSSNVHEPGFICGAVSKLDY 232
Db 132 DLY-QHRRG--LSFDDEWENDSPLDVPLEDMVYEMHVRGFTRESKIEPPTGYQVAEKLD 188
Oy 233 AVSKLDYKELGVNCLMPCHEFNELEYS-----TSSKMFVGYSTINEFSPMRYTSGIKNC 287
Db 189 ILKIRIYDGLVNTIELMFIPEFDEPHSRYPETGEFLVNTWGYSTVNFAPAKAGYAA 248
Oy 288 GGKRCGRDAINEKTEFVREAHKRGIEVILDVYFNHTAEGNENGPILSEFGVNDITYYML 347
Db 249 TG--KFGMG--IDBLKMLVKEHLKVGISVILDVYFNHTAEGNENGPILSEFGVNDITYYML 305
Oy 348 APGGEFYNYSGCGTENCNHPVYRQFTYDCLRYVWTEMHVDFRFDLASTINTRGSSSLMD 407
Db 306 TPBGYTFNFGTNTLNCNPIVRCVGLDCLRTWTAEFHIDGFRFDLASTIGR-----DP 360
Oy 408 VNYGARIIEGDMITGTPLTPIIDMISNDPIIGVKLLAEAMDAGGLYQVGOFPMMNVM 467
Db 361 -----WGYPPLANPLETLADPLIASKLLAEAMDAGGLYQVGOFPMMNVM 406
Oy 468 WSEMGKYRDIIVROFIGTDFAGGAFACLCGSPHLVYQAGRRPWHISINFCVCAHDEFTL 527
Db 407 WAEWNGKYRDIIVROFIGTDFAGGAFACLCGSPHLVYQAGRRPWHISINFCVCAHDEFTL 466
Oy 528 DLYTNKKYNLPNGENNRDGENHNLISWNGEYSELHRCCLMTRFKRECEGLDEDEPTAKRL 587
Db 467 DLVAYNGKHNYAENGNDANDNYSWNGCEBPTDNPDLRLAROMRAIALIVSOG 526
Oy 588 VPMFYNGDEGHTKGGNNNTYCHDSYVNYRPMWKKEQYSELHRCCLMTRFKRECEGLDE 639
Db 527 VPMFYNGDEGHTKGGNNNTYCHDSYVNYRPMWKKEQYSELHRCCLMTRFKRECEGLDE 586
Oy 640 -----KECEGLDEDEPTAKRLQMHGHOGRKPDMSSENSRFVAFSKMDE-----ROGEIY 688
Db 587 SEHFQNDYLGV-GFPD---ISWGVKPMHADSSADSRVLAFLMCGNAHAGRKNDQIY 642
Oy 689 VAFNTSHLPAVELEPERAGRWEPVDTGKPARVYELTDLDRALTIH 723
Db 643 VAMNMHYESLMEFLPARPVGTWVHVFANTGAOPPED 678

```

## RESULT

5

isoamylase-like protein [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C:Accession: B85091

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold ;

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.

A:Reference number: B85091

A:Accession: B85091

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-702 <STO>

A:Cross-references: GB:NC\_001268; NID:g7267544; PIDN:CAB78026.L; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g09020  
A:Map position: 4  
C:Superfamily: glyx protein

Query Match 36.5%; Score 1538.5; DB 2; Length 702;  
Best Local Similarity 44.0%; Pred. No. 2.8e-110;  
Matches 325; Conservative 110; Mismatches 222; Indels 81; Gaps 19;  
56 RVLGAPALGATAGLVNFAVYSGATAAALCLF-----PEDKADRYTEVP 106  
8 KYSSGSEVSLGVQVQKGNFALFSONATSVTCLISQRTTSSGKDDTDDGMIPLV 66  
107 LDPMLNRTGNVWHVIEG-ELHNMXYRGDTFAPRCHYLDSNVVDPYAKAVISRG 165  
67 LDPVNVKTDGTHICVEDLPLNNVLYGYRVDGEMQGRHFRSTILLDPYAKLVKGS 126  
166 EYGVPARGNKCPOMAGMTPRYSTDMGCDLPL-RTPOKDLVYEMHLLRGFTKHSSNV 224  
127 SFG---DSQKFAQFYGTDFESSPDMGDDYKFPNIPKEDLVYEMNVAFTADESSGM 183  
225 EHP--GTFGAVSKLDYLKELGVNCTELMPCHEFNELEYSTSS---KNMFGYSTIN 278  
184 DPAIGSYIGFTLEKIPHLQDLGINAVELLEVEFEDELELORSNPRDHNVWGSTVNF 243  
279 FSPMTRYTSG---GIKNCGRDAINEKTFVREAHKRGIE-----VILDVVFNTAE 326  
244 FAPMSYVAGEGDPK----ASKERKEMVKALHSAGIEKYSYKFSQVILDVVYHNTNE 298  
327 GNENGP-IISFKGVNDVTYMLAPKGEFYNSCGNTFNCNNHYVQFIVDCIRYVTEM 365  
299 ADDKYTYTTFRQIDMKRYVYMLDPNNQLNFSGCGNTLNCNHPVYMLILDSLRHWTEY 358  
386 HDVGFPEFDLASITRGSSSLMDPVNVYCAPLEGDMITGTPLVTPPLIDMISNPIIGVK 445  
359 HDGFFFDLASVYCR-----IDGSLSPAPLIRAIKDSVLSRCK 399  
446 LLAEMADAGLYGVGFPHNMVSEWNGKRYDIVROFIKGTDFAGGFAECLGSPHLYQ 505  
400 ILAEPDCCGLYLVGKFRPMDBMAENNGMYRDVRRIRKDSGKSGFAIRVSSSLYQ 459  
506 AGGRKAWHSHNFYCAHDGFTLADLVYNNKYNLPCNGENNDGEGHNLNMGEGEGERL 565  
460 VNQRKRYHGNFYIADGFTLRDLVSYNFKNHANEKGEGDCNDMSNMGEGEGEDA 519  
566 SVKRLRKRMNFVYCLMVSGVPMFYMGDEYHGTGNNNTYCHDSYVYFRMRDEKQY 625  
520 HIKSLTRQKNFHLALMISQGTPMMLMGDEYHGTIRGNNSYGHDTSLNPFQKELDAK 579  
626 SELH-RFCCLMTRFKRECEGLADEFTAKRLQMHQHPGKPDV-SENSRFAVSMKDER 683  
580 KQNHFFFESEVLIKFRSHNHVHLKHENFLTOGEIITWH-----EDMNDNSESFLAFTLDGI 634  
684 QG-EIYVAENTS--HLPAVELRERAGRWEPVVDGTGRAPYDFLJLDDRLATTHQFS 740  
635 GGDIVYAFNAHDYFKALIPDP-PEKQWRVADTILSPDDFVRGVGVAVDT----- 688  
741 HFLYSNLVPLMYSVIL 758  
689 -----YVAPFSSIL 699

RESULT 6  
S74546  
glycogen operon protein (EC 3.2.1.-) glgX-1 - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: protein slr0237  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S74546  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ye  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S74546  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-746 <KAN>  
A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BA16698.1; PID:g16;  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: glgX-1  
A:Start codon: GTG  
C:Superfamily: glyx protein  
C:Keywords: glycosidase; hydrolase

Query Match 33.0%; Score 1388; DB 2; Length 746;  
Best Local Similarity 42.0%; Pred. No. 1.2e-98;  
Matches 305; Conservative 86; Mismatches 251; Indels 84; Gaps 17;  
60 GMPAPLGATAGLVNFAVYSGATAAALCLF--TPEDLKADRYTEVPDPLMNRGNV 118  
17 GQTFPLGATYVADGVNFCLESKAERTLLDFDRNPAPARTIE---LIRGRRTTYW 73  
119 HVEIEGLHNLXYRGDTFAPRCHYLDSNVVDPYAKAVISRGY---GVPARGN 175  
74 HVEVKGKLAGQVAYRVGDEHEBEKGRFPDVKLVYKAKAIYGRKIDKAMALADN 133  
176 CWPQMGMITPLPYSTFPMEDDLPLRYPKDLVYEMHLLRFTKHSNVH--PGTFIGA 233  
134 C-AQALRSVVDVSYVDMEDHAPRPYASITELVAGPTRPNGLSBNKRGYAGL 192  
234 VSKLDYLKELGVNCTELMPCHEFNELEYSTSSKMNFGYSTINFSPMTRYTSGIKNC 293  
193 IEKIPYKELGTRAVELLPHYHFDPE--AQPGITNWGSTIGFFAPHGYSA---DD 247  
294 GRDAINEKTFVREAHKRGIEVILDVFNHTAEGENGPILSEKGVNDNTYYML-APKGE 352  
248 PLEVDEDFRDMVVALKRAKIEVILDVFNHTAEGENKPTLSFRGINRTYYILDEDKSS 307  
353 FNYVSGGNTFNCNHPVRYVDCIRYVTEMHVDFFRDLASIMRGSLSMPVNVYQ 412  
308 YSNVSGGNSVKNHPVVGGLILDSLRVWSEMHVDFRFDLASVLRDRK-----G 359  
413 APIEGDMITGTPLVTPPLIDMISNPIIGVKLIAEAMPAGGLYGVGFPHNMV-SEW 471  
360 VPLHGSATIAN-----ITAIIESDPILAGTKLIAEAMDAAGLYSVKRYELADNFAEM 413  
472 NGKRYDIVROFIKGTDFAGGFAECLGSPHLYQAGGRKPMHSINFVCAHDGFTLADVT 531  
414 NGPRDVRVRFVAGDGNAPALASRLSGSPDIYRQOTDIRNSINFVTCIDGFTLVLD 473  
532 YNKRYNLPNEENRDNHNLNMGEGEGEFARLSVKRLRKQRMNFVCLMVSGVPMF 591  
474 YNEKHANNEKKNKDNDFSNMGVGETDPIKQLRLQIKNLTILFFSGQPM 533  
592 YMGDEYHGTGKGNNTYCHDSYVYFRMDKKEQSEHLRCK----- 632  
534 LMGDVPGRTQGGNNGCCODNELSMFMSGEHSDETHLRGLTALQSLPLEBSLL 593  
633 ---CLMTKFRKE--CEGLGLE-DFPTARLQ-----WHG 660  
594 PVLNFAIPLRPDENDGSEGGODEEVPYRPLPVDEBALREMYLKATKDLPGSCVYWHG 653  
661 HQGKPRPMSNSRVAFAVSMKDERGEL-YVAFNTSHLPAYVELEP-RAGRWEPVVDGTG 718  
654 VKLAQPDWSTYSHLAVTLFHPGAETILHLIFNAWYELNLEPRLNDGLIMRLVDVYL 713  
719 PARYDF 724  
714 PTFPLDF 719

RESULT 7  
B86539  
glycogen hydrolase (imported) - Chlamydomonas reinhardtii (strain J138)  
C:Species: Chlamydomonas reinhardtii, Chlamydomonas reinhardtii  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: B86539  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: B86539  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:BA000008; NID:98978760; PID:BA098596.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: glgX  
C:Superfamily: glgX protease

Query Match 31.0%; Score 1305.5; DB 2; Length 664;

Best Local Similarity 42.6%; Pred. No. 2.3e-92; Mismatches 231; Indels 57; Gaps 16;

Matches 279; Conservative 86; Mismatches 231; Indels 57; Gaps 16;

61 MPAPLGATATA-GGVNFAVSGGATTAALCLFTEPDLKADRYTE--EYPLDPLMNRGTNV 117  
10 VPLPLGASKISPNRYRFAALASQATEVIAL-----TDENSEVIEVPLDTHRTGAI 62  
118 WHVFEGELHMLGYRFDGTFAPHCCHYLDVSNVNVVPAKAVISREGVCPARGNN- 175  
63 WHIEGIDSDSSAFRHPG-KKH-GMOYSFKEYLADPYAKNHSPQSFRRKKGQDYA 120  
176 -CMPQAGMPLPYSTFPMWGDPLRYPOKDLVYEMHRLGFTKHDSSNVNHEPFTIGAV 234  
121 FCY-----LKEEPFMWGDQPLHLKPEEMITTEMHVRSTQSSSRVHAPGFTIGAI 172  
235 SKLDYKELGVNCLMPCHEPNELEYSTSSKM---NEWGSTINEFSPMTRVTSIGI 290  
173 EKIDHLKLGINAVELLPIFEDETAFHFRNSKFPYLCNMGVAPLNFESCRAYAAVS- 231  
291 KNGCRDAINEKFTVREAHKRGIEVILDVVFNHTAEENENGPILSFQVNDTTYMLAPK 350  
232 DPCAPS--REFKTLVKTLDHOGIEVILDVVFNHT--GLOGTCSLPMIDTPSYIIDAQ 286  
351 GEFYNSGCGTFCNCPNRPVROFTVDCLRVWTEHVDGFFRDLASIMTRSSLMDDPVNV 410  
287 GHFTNYSGCCNTLNTNAPFTOWILDLRYVEEMHVDGFFRDLASVFSRGS----- 339  
411 YGAPIDGMITTGPLVTPPLIDMISNDPILGVKLIAEADAGLXOVGFPPHNV-WS 469  
340 -----GSPLOFAVLEAISFDPLASTKIIEAPDAGGLXOVGFPLSPRMS 387  
470 EWMNGKRDIVROFTIKGTDFAGFAECLCGSPHLYOAGGRPMHSINVCACHDGTIADL 529  
388 EWMNGPRDYNKAFELNGDONLIGTFASRISGQDIYPHG--SPTNSINVSCHDGTICDT 445  
530 VTYNKKYVLPNGENNRGENHNLNMGCEGEFARLSVRLRKROMRNFVCLAVSOGVP 589  
446 VTYNKKHNEANGENNRGTSTANYSYNGTEGKTEDPGLLEVRERQLRNFPLTLAVSOGIP 505  
590 MFTYNGDEYGHKGGNNNTYCHDSYVNYFRMDKROYSELARFCCIMTKFRCEGLGLED 649  
506 MIOGGBEYATTAAGNNRNALDLSNANFYLDQTLAKPTLHAFGLCDLAFKRYKTYTLNRG 565  
650 FPTAKRLQWGHOPGKRPDSSENSRYVAFSKDERQETIYVAFNTSHLPAVELDE 704  
566 FLNKEISVWDAGNPMTRPGN-FLAFKTKSPK-AHYVYAFHVGADQDLATLPRK 618  
RESULT 8

B72084

glycosyl hydrolase family protein CP0367 (imported) - Chlamydomonas reinhardtii (strain J138)  
C:Species: Chlamydomonas reinhardtii, Chlamydomonas reinhardtii  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: B72084; C81586  
R:Kolman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Ollinger, L.; Grimwood,  
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydomonas reinhardtii and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: B72084  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <ARN>  
A:Cross-references: GB:AE001623; GB:AE001363; NID:94376662; PID:AA018532.1; PID:94;  
A:Experimental source: strain CML029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gail, S.R.; Heidelberg, J.F.; White, O.; Hici,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, C.; Salz,  
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydomonas reinhardtii and Chlamydomonas trachomatis AR  
A:Reference number: A81500; MUID:20150255  
A:Accession: C81586  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <RNA>  
A:Cross-references: GB:AE002198; GB:AE002161; NID:97189279; PID:AAF38216.1; PID:97;  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: glgX; CP0367  
C:Superfamily: glgX protease

Query Match 31.0%; Score 1305.5; DB 2; Length 664;

Best Local Similarity 42.6%; Pred. No. 2.3e-92; Mismatches 231; Indels 57; Gaps 16;

Matches 279; Conservative 86; Mismatches 231; Indels 57; Gaps 16;

61 MPAPLGATATA-GGVNFAVSGGATTAALCLFTEPDLKADRYTE--EYPLDPLMNRGTNV 117  
10 VPLPLGASKISPNRYRFAALASQATEVIAL-----TDENSEVIEVPLDTHRTGAI 62  
118 WHVFEGELHMLGYRFDGTFAPHCCHYLDVSNVNVVPAKAVISREGVCPARGNN- 175  
63 WHIEGIDSDSSAFRHPG-KKH-GMOYSFKEYLADPYAKNHSPQSFRRKKGQDYA 120  
176 -CMPQAGMPLPYSTFPMWGDPLRYPOKDLVYEMHRLGFTKHDSSNVNHEPFTIGAV 234  
121 FCY-----LKEEPFMWGDQPLHLKPEEMITTEMHVRSTQSSSRVHAPGFTIGAI 172  
235 SKLDYKELGVNCLMPCHEPNELEYSTSSKM---NEWGSTINEFSPMTRVTSIGI 290  
173 EKIDHLKLGINAVELLPIFEDETAFHFRNSKFPYLCNMGVAPLNFESCRAYAAVS- 231  
291 KNGCRDAINEKFTVREAHKRGIEVILDVVFNHTAEENENGPILSFQVNDTTYMLAPK 350  
232 DPCAPS--REFKTLVKTLDHOGIEVILDVVFNHT--GLOGTCSLPMIDTPSYIIDAQ 286  
351 GEFYNSGCGTFCNCPNRPVROFTVDCLRVWTEHVDGFFRDLASIMTRSSLMDDPVNV 410  
287 GHFTNYSGCCNTLNTNAPFTOWILDLRYVEEMHVDGFFRDLASVFSRGS----- 339  
411 YGAPIDGMITTGPLVTPPLIDMISNDPILGVKLIAEADAGLXOVGFPPHNV-WS 469  
340 -----GSPLOFAVLEAISFDPLASTKIIEAPDAGGLXOVGFPLSPRMS 387  
470 EWMNGKRDIVROFTIKGTDFAGFAECLCGSPHLYOAGGRPMHSINVCACHDGTIADL 529  
388 EWMNGPRDYNKAFELNGDONLIGTFASRISGQDIYPHG--SPTNSINVSCHDGTICDT 445  
530 VTYNKKYVLPNGENNRGENHNLNMGCEGEFARLSVRLRKROMRNFVCLAVSOGVP 589  
446 VTYNKKHNEANGENNRGTSTANYSYNGTEGKTEDPGLLEVRERQLRNFPLTLAVSOGIP 505  
590 MFTYNGDEYGHKGGNNNTYCHDSYVNYFRMDKROYSELARFCCIMTKFRCEGLGLED 649  
RESULT 9

Db 506 MIOGDEVAHTAECNNRMALDSNANYFLMDQLTAKPTLMHFLCDLAFKKKYLTFNNG 565  
QY 650 FPAKRLQWGHQCKPOMSENSREVAFSMKDEROGEIYAANTSHLPVVELPE 704  
Db 566 FLNKSISWAMVAMGNPMTWRPQN-FLAFKIKSPK-AHYVAHVGAQDQDLATLPK 618  
RESULT 9  
E71565  
probable glycosylase (debranching) - Chlamydia trachomatis (serotype D, strain UW  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: E71565  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac  
A:Reference number: A71570; MUID:99000809  
A:Accession: E71565  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-666 <ARN>  
A:Cross-references: GB:AE001278; GB:AE001273; NID:g3328423; PIDN:AA67632.1; PID:g332843  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: glxX  
C:Superfamily: glyx protein  
Query Match 30.9%; Score 1300; DB 2; Length 666;  
Best Local Similarity 39.8%; Pred. No. 6-2e-92;  
Matches 289; Conservative 102; Mismatches 241; Indels 94; Gaps 20;  
QY 57 VLAGNPAPLATAL-AGGVNFAVYSGGATAALCLFTEPDLKADRYTE--EYPLDPLMNR 113  
Db 6 VRSTPLPLGAKKLSADRYRFLSSQAQOQVTLVLDLP-----LSEIHEIPLSSTDHR 58  
QY 114 TGNVWVHVEIEGLHMLKCYRFDG-----FAPHCGLDVSNNVVDYAKAVISRGEX 167  
Db 59 TGAIHIEIAGISSEMSYAYKLRGDLSSQKPA-----TDSYADYSKNIYSPOLE 110  
QY 168 GVPARGNCPOMAGMIPLPYSTFMEGDLPLRYPOKDLVIEYEMHGRGTTKHDSNVNEHP 227  
Db 111 GSPKOEK-----YAFSYLKHEDFMEGDTPLHPKRENFIEYEMHGRSTTRODSQSV 165  
QY 228 GTFICAVSLDYKLELVNCIELMPCHEFNELEYSTSSKM---NFWGYSTINFEPSMT 283  
Db 166 GTFLIIEIKIDHLKOLGVNAVELLPFEFDETVHPKNOFPHLCYMGYSVNFECPSR 225  
QY 284 RYTSGLNCGDAINEKTFVREAHKRGIEVILDVFNHTAEGNNGPILSKGVNDNT 343  
Db 226 RYTVGA-DPCA--PARFEFTLVKALHRAIEVILDVFNHTGEGTSCPL--PWIDLES 279  
QY 344 YVMLAPKEEFYNYSGCNTFNCNHPVROFIVDCLRYVTTEMHVDFRFDLASIMTRGSS 403  
Db 280 YVMDHODLDMFSCGNTVNTPTTLKIIDALRYVQEMHVDGFRFDLASVSR--- 336  
QY 404 LMDPVNVYGAPIEGDMITTTGTPPLVTPPLDMSNDPILGVKLIABWADAGGLYQVGP 463  
Db 337 --D-----OGVPLPLPLPILQAISSDSIISLSEKLIABPMDAGGLYQGHFP 380  
QY 464 HNVN-WSENGKRYRDIVROFIKGTGCFAGGFAECLOGSPHLVQAGRKPMHSINFCAND 522  
Db 381 S1STMSSENGCYRDHVKFLNGDAHQVSSFASRISGSDIYPNG--KPTNSINNYCSHD 438  
QY 523 GFTLADLVYNNKYNLPNENNRDGENHNLNWNNGEGEFARLSVRLKRRQMRNFVCL 582  
Db 439 GFTLIDVYANNKHNENNEYNRDGSANYSYNFCGEGETPTPTICALERQKKNFFAL 498  
QY 583 NVSGVPMEFYMGDEYGHTKGANNNTYCHDSYVYFRWDRKKEOYSELARPCCLMTKFRKEC 642  
Db 499 FLSCGIPMIOGDEYGHYTAAGNNHWCIDTKINYFLMDRLAEKELFSFLCYIALRKAY 558  
QY 643 ECLGLEDFPTAKRLQWGHQCKPOMSENSREVAFSMKDEROGEIYAANTSHLPVVEL 702

Db 559 TELFNNSFLSEDDPTITWLNKSGSPREWGAD-HYLAFLKL-HLNSLFLVAFYSGNERIEISL 616  
QY 703 PE--RAGRMEEVVD--TKRPAYDPLTDLPDRALTIQSFHFLXSNLYPMHS---YSS 755  
Db 617 PPRKREHLAYEKIVDSTG-----FFSQILSPKLSLEPYSS 652  
QY 756 VILVLR 761  
Db 653 LVAISR 658  
RESULT 10  
B75540  
glycosylase operon protein Glx - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: B75540  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75540  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-720 <WHI>  
A:Cross-references: GB:AE001888; GB:AE000513; NID:g6457936; PIDN:AAF09848.1; PID:g64;  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0264  
A:Map position: 1  
C:Superfamily: glyx protein  
Query Match 30.1%; Score 1268; DB 2; Length 720;  
Best Local Similarity 39.8%; Pred. No. 2e-89;  
Matches 282; Conservative 92; Mismatches 241; Indels 94; Gaps 17;  
QY 60 GNPAPLATALAGVNFAYVSGGATAALCLFTEPDLKADRYTEYVPLDPLMNRGNNVH 119  
Db 11 GSPFPLGATWMDGNTGTFALYSENAIVGELCLFPAEHEHET-----RFLTGOTAFVNH 62  
QY 120 VETIEGLHNNKLYGRFDGTFAPHCGLDVSNNVVDYAKAV-----ISNGEYVAPARGN 174  
Db 63 GYLPRIQOPQGRYGVRAHGEVAPAEKGRFNPNVVLDPBYAKALDGTGEQFDGFGVGVYAGGE 122  
QY 175 NCPKPK-----AGMIPLPYSTFMEGDLPLRYPOKDLVIEYEMHGRFT-KHDSNV 224  
Db 123 DD--SOMEDEEORGAPGLVVDP--MFNMGDQKRGIPFHOSVIEAHVKGLTMTHPDVE 179  
QY 225 EHPGTIGAVSK--LDYLKELGVNCIELMPCHEFNELEYSTSSKKNFMGYSTINFEPSM 282  
Db 180 ELRGYTAGVATPAIIDLNDLGLITAEPLRVHGHVDDPLDGLGNTYWKYSTLNFAPD 239  
QY 283 TRYTSGLNCGDAINEKTFVREAHKRGIEVILDVFNHTAEGNNGPILSKGVNDNT 342  
Db 240 VRSASABARKGNPAGVAPFEKFNVRALHDAGIEVILDVFNHTAEGNNGPILSKGVNDNT 299  
QY 343 TYVMLAPKE--FYNSGCGNTFNCNHPVROFIVDCLRYVTTEMHVDFRFDLASIMTR 400  
Db 300 TYRILVADDOREFYDTGTGNSLVNRPOTLQIMOSLRVWTEMHVDFRFDLASLAR 359  
QY 401 GSSLMPPVNVYGAPIEGDMITTTGTPPLVTPPLDMSNDPILGVKLIABWADAGGLYQV 459  
Db 360 GLHEVQLSGF-----FTIHODPLISQVKKLIABWMDGEGGYOV 359  
QY 460 GQFPHNVNVEGKRYRDIVROFIKGTGCFAGGFAECLOGSPHLVQAGRKPMHSINFC 519  
Db 400 GNP--VNNAEWNGIYRDMRSFWMKGGGLASEIGYITGSSDLYEFGNKKRPASINFTY 457  
QY 520 AHDGFTLADLVYNNKYNLPNENNRDGENHNLNWNNGEGEFARLSVRLKRRQMRNF 579





```

Oy 397 IMTRGSSLMDDPVNYGAPIEGDMITTGRLVPLPLDMSNDPILGKYKLAEMDAG-G 455
Db 366 TLAEHLHDVDRLSAF-----FDLVQODPVVSOYKFLIEPMDVGSG 405
Oy 456 LYQVGFPPHMMVWSEMNKRYDIYRQFTKGTDFAGFAECISGPHLYQAGRRPMHSI 515
Db 406 GYQGNFPP--GLMTPEMNKRYDVTADYWRGEPATIGEFASRLTSSDLYEAGRRPSASI 463
Oy 516 NFVCAHDEFTLADLVYTKKTKNLPGENNRDGENHLSMNGCEGEFARLSYKRLKROM 575
Db 464 NFVTAHDEFTLNDLVSYNDKHNEANGENNRDGEYSNRSCNCGEPTDDPDLALRAROM 523
Oy 576 RNFVCLMVSQGVPRFVYMGDEGHTKGGNNNTYCHDSYVNYRMRKKEGYSLSHRECCLM 635
Db 524 RNMATLWVSQGTPIAHGDEIGRTQYGNNNNYCODSELISMWDSLVQDNADLAFARKA 583
Oy 636 TKFRKE-----CEGLGEDFPTAKRLQW---HGHPQKPRDSEN--SRFAESMKDE 682
Db 584 TLLKKNHKVFRRRRFEDEPIRSGDEVADIAMLTPSGRMTHEDEMGRRGRCVAVAFPLNGE 643
Oy 683 -----RQGE-----IYAEFTSHLPAVELLP-ERAGRRMEPVVDTKRP 719
Db 644 AITAPDARGERVDDSFLLCFNAHDHVEFVMPHDGYAQOMTGELDTNDP 693

```

## RESULT 13

```

glycogen operon protein glxg (EC 3.2.1.-) - Sulfolobus solfataricus
S73088
N:Alternate names: protein c0621
C:Species: Sulfolobus solfataricus
C>Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S73088
R:Seisen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal genome: I
A:Reference number: S73076; MUID:97055432
A:Accession: S73088
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-718 <SEN>
A:Cross-references: EMBL:Y08256; NID:q1707679; PIDN:CAA69504.1; PID:q1707700
A:Experimental source: strain P2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Keywords: glycosidase; hydrolase

```

Query Match 28.8%; Score 1211; DB 2; Length 718;

Best Local Similarity 38.2%; Pred. No. 5e-85; Indels 106; Gaps 19;

```

Matches 274; Conservative 101; Mismatches 236; Indels 106; Gaps 19;
Oy 60 GMPAPLGTATAA--GGVNFAYVSGGATAAALCLFTPEDLKAADRYTEVPLDPLMNRGNV 117
Db 15 GDRPVLGSNMIEDDDGVNPSLFSENAKEKELLSTLNKYKKEITLV-----KNTGDI 69
Oy 118 WHVFIEGELHNMKLYGRPDGTFAPHCGLVDSNVVVDYAKAV-----161
Db 70 WHVFEPGLRPGOLYAYRVYGPYKPELGLRFNPNKVLIDPYAKAINGSVIWMNDAVEGKIG 129
Oy 162 -----ISRGEXYVAPARNNCMPQAGMIFLPYSFFDMGSD---LPLKYPQKDIYI 208
Db 130 DONODLVYDERDSGEY-VP-----KSVINPY--FEWDDDEDFIKGKKVPLKDTYI 176
Oy 209 YEMHLRGFTKHDSSNVEH--PGTFIGAVSK--LDYLKELGVNCLIEMLPCHFEFNELESTSS 265
Db 177 YEVNHNKGTTLRLDLPENIRGTVEGLASBQMTSYLKDLGITVELMPLPVPHFDQRLITXK 236
Oy 266 SKMNFVSGTINFSFPMTRYSSTGKICNGRDAINEKTFVREAHKRGIEVILDVVFNHTA 325
Db 237 GLTNWGYDPIINFESPECHYSTSG--CLGQGVLSFKKMYNELHNAIGIEVILDVVYNHTA 293
Oy 326 EGNENGPILTSFGVNTTYMLAPKGEF--NYSGGCNFNCHNHPVRRFYTDCLRYWYT 383

```

```

Db 294 EGNHILGPTLSFGINDTAYATMLQPOMKRYTLDFTGTGNLTLSHPRIQMDVDSLYRWYT 353
Oy 384 EKHVDSFRFDLASITRGSSLMDDPVNYGAPIEGDMITTGRLVPLPLDMSNDPILG 443
Db 354 EKHVDSFRFDLAALAR-----ELYSV-----NMLNT-----PFIALQODPILSO 393
Oy 444 VKLIAEMDAG--GLVOVGFPMMHMMWSEMNKRYDIYRQFTKGTDFAGFAECISGPH 502
Db 394 VKLIAEMDAG--GLVOVGFPMMHMMWSEMNKRYDIYRQFTKGTDFAGFAECISGPH 451
Oy 503 LYQAGRRKPMHSHINFCADGFTLADLVYTKKTKNLPGENNRDGENHLSMNGCEGEF 562
Db 452 IYLGNNKTPFAINVTYSHDFTLEDVSYNKHNAKPNFNDGNNENYSHNCGADEPT 511
Oy 563 ARLSYKRLKROMRNEFVCLMVSQGVPRFVYMGDEGHTKGGNNNTYCHDSYVNYRMRK 622
Db 512 NQNNVYICREKOKRNMIMILLVSOGFTMLGDELSRTORGNNAPACQDNEITWFMND 571
Oy 623 EQLSELHRECCMTKFRKCEGLGLEDFPTAKRL-----QMHGHPQKPRMSEN 671
Db 572 EKRSKFLFVKKMIQIYRAHPAFRRERYQKKLFGMPLKDVTFYTLLEGREVEDEKTS 631
Oy 672 SRFAVSMKDEROGEI-----YAEFTSHLPAVELLP-ERAGRRMEPVVDTKRP 716
Db 632 TQLVIFVLEGSYMDIEMNGERLADDSFLILNANNNVKKVPRK---GKMLVYSS 685

```

## RESULT 14

```

C83375
probable glycosyl hydrolase PA2160 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 24-May-2001
C:Accession: C83375
R:Stover, C.K.; Pham, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.;
Loy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p
A:Reference number: A82950; MUID:20437337
A:Accession: C83375
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-716 <STO>
A:Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AA05548.1; GSPDB:G
A:Experimental source: strain PA01
A:Gene: PA2160
C:Superfamily: glyx protein

```

Query Match 27.3%; Score 1149.5; DB 2; Length 716;

Best Local Similarity 36.4%; Pred. No. 2.7e-80; Indels 117; Gaps 24;

```

Matches 275; Conservative 103; Mismatches 260; Indels 117; Gaps 24;
Oy 56 RYLACMPAPLGTATAGVNFAYVSGGATAAALCLFTPEDLKAADRYTEVPLDPLMNRG 115
Db 11 RISEGHFPLGATWDLGVNFAFLFSAHAKRYELCLP---DAREKIEIRLELP---EYTD 64
Oy 116 NMVHVFIEGELHNMKLYGRPDGTFAPHCGLVDSNVVVDYAKAVISGEYVAPRGN 175
Db 65 EIMHGYLPDHPAQITGYVHGHYEPDAGHREPNKLLDPTAKOLVGRMSEALRGYT 124
Oy 176 CWPQAGM-----IP-LPYS-----TFDWEGBDLPRLYPOKDLVYEMHLRGFTKHDSS 222
Db 125 IGSADADLSFDERDSAPFVPRKSVIDPAPFTWAEPRVVRPMDRTVYEALRL-----178
Oy 223 NVEBP-----GTFIGAVSK--LDYLKELGVNCLIEMLPCHFEFNELESTSSKKNFMGY 273
Db 179 SMRHPOVRAVAVGTFAGLNADLAHRIIGLVYSVELLPHGCVDDKLLLENQMSNWTG 238
Oy 274 STINFSFPMTRYSSTGKICNGRDAINEKTFVREAHKRGIEVILDVVFNHTAGNNGPI 333

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 12:11:11 ; Search time 25.26 Seconds

(without alignments)  
1171.090 Million cell updates/sec

Title: US-09-674-817A-2

Perfect score: 4212

Sequence: 1 SGAPAPRLRRMPNATAGKV.....SNLYPMSTSYIVLRLPDV 764

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325	29.1	721	1	GLGX_MYCTU
2	1014	22.1	659	1	GLGX_HAEIN
3	1001	23.8	657	1	GLGX_ECOLI
4	810	19.2	777	1	ISOA_FASP
5	803.5	19.1	776	1	ISOA_PESP
6	798.5	19.0	776	1	ISOA_PSEAV
7	523	12.4	843	1	PURA_THEMEA
8	310.5	7.4	562	1	OLIG_BACTR
9	310	7.4	598	1	TREZ_ARTSP
10	283.5	6.7	596	1	TREZ_RHISP
11	278	6.6	575	1	TREZ_ARTRM
12	277.5	6.6	589	1	TREZ_BREHE
13	272	6.5	580	1	TREZ_MYCTU
14	265	6.3	1090	1	PURA_KLEPN
15	255	6.1	585	1	NEPU_THEMEA
16	249	5.9	588	1	NEPU_THEMEA
17	247.5	5.9	639	1	GLGB_BACST
18	247	5.9	1096	1	GLGB_KLEAE
19	246.5	5.9	591	1	CDAS_BACSH
20	246.5	5.9	1481	1	APU_THEMEA
21	245.5	5.8	728	1	GLGB_ECOLI
22	243.5	5.8	666	1	GLGB_BACCL
23	241.5	5.7	558	1	OLIG_BACCE
24	241	5.7	702	1	GLGB_HUMAN
25	241	5.7	799	1	GLGB_MAIZE
26	238.5	5.7	586	1	AWYM_BACAD
27	236	5.6	561	1	OLIG_BACAD
28	234.5	5.6	639	1	GLGB_BUTET
29	231	5.5	535	1	DEXB_STRPN
30	228.5	5.4	773	1	GLGB_SYNPF
31	228	5.4	498	1	AMY3_DICTH
32	228	5.4	731	1	GLGB_MYCTU
33	221.5	5.3	1473	1	APU_THEMEA

34	220	5.2	770	1	GLGB_SYNPF
35	219.5	5.2	567	1	MAU1_APIME
36	218.5	5.2	574	1	CDAS_THEMEA
37	215.5	5.1	573	1	TRES_PIMSR
38	214.5	5.1	627	1	GLGB_BACSU
39	212.5	5.0	581	1	MAXS_YEAST
40	208	4.9	764	1	GLGB_STRAU
41	205.5	4.9	1279	1	APU_THEMEA
42	201	4.8	562	1	AMY2_DICTH
43	200	4.7	561	1	TREC_BACSU
44	200	4.7	734	1	GLGB_AGRU
45	197	4.7	1196	1	AMYB_PABPO

## ALIGNMENTS

RESULT 1  
ID GLGX\_MYCTU STANDARD: PRT: 721 AA.  
AC Q10767;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycogen operon protein glgx homolog (EC 3.2.1.-).  
GN GLGX OR RV1564C OR MT1615 OR MTCY48.01.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,  
RA Bishai W., Ullrich T., Weidman J., Khoult H., Gill J., Mikula A.,  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC - SIMILARITY: BEGONS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY. STRONG, TO E.COLI AND  
CC H.INFLUENZAE GLGX.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
DR EMBL: 274020; CA98327.1; -  
DR EMBL: AE007027; AAK45882.1; -  
DR HSSP: P10342; 1BF2.  
DR TIGR: MT1615; -  
DR TubercuList: RV1564C; -

DR InterPro: IPR000461: Alpha.amylase.  
 DR InterPro: IPR004193: Isoamylase\_N.  
 DR Pfam: PF00128: alpha-amylase; 1.  
 DR Pfam: PF02922: Isoamylase\_N; 1.  
 KW Hydrolase; Glycosidase; Complete proteome.  
 FT ACT SITE 362 BY SIMILARITY.  
 FT ACT SITE 470 BY SIMILARITY.  
 SQ SEQUENCE 721 AA; 81081 MW; 5616457CDCC87D1 CRC64;

Query Match 29.1%; Score 1225; DB 1: Length 721;  
 Best Local Similarity 39.4%; Pred. No. 1.2e-85;  
 Matches 280; Conservative 85; Mismatches 263; Indels 82; Gaps 19;

57 VLAGMAPLGTALAGVNAVYSGGATAAALCLFTPEDLKADRYTEVPDPLMNRGN 116  
 19 VMPGNAPPLGATYDGAGTFSLEIAEKVELCL-----IDEDGESIRPLEDEV---DGY 70  
 117 VMHVFIEGELHNLGYRFEDGFAPHCCHYLDVSNVVDPAKAVISRGEG----- 168  
 71 VMHAYLPNITTPGGRGFVHGRFPRAAGHRCDDPSKLLDPPGKSHGDFTEGQALYSTDV 130  
 169 -----VPARGNCWPMAGMIPPLPYSTFPMESDPLRYPOKDLVIYEMHRCFTK-HD 220  
 131 NADPDPSTPPAVDSLGHTMTSVINPF--FDMAIDRSPTREHETVIEAHKMGMTQHP 188  
 221 SSNEHGTFIGAVSK--LDYLKELGVNCEILMPCHEFNELEYSTSSKMNFGYSTINF 278  
 189 SIPELNGTYAGLHAPVLIHNLNLTAVELMHPQFLHDSRLDLGLRNYMGVYNTGCF 248  
 279 FSPRTRTSGIKKCGRAINEFTFVREAKRGIEVILDVVFENFAGNENGPILSPFG 338  
 249 FAPRHQYAS--TRQAG-SAVAEFTKTVRSLEHAGIEVILDVVYNTAGCNHGLPITNFG 305  
 339 VDNITYMLARKG-EFY-NYSGCGNTFNCNHPVYVQFIVDCRYVTEMHVDFGFFDLAS 396  
 306 IDNATYRLMDHDLRFYKDFGTGNSLMARHNTQLLMDLSRYVIEKHNVDGFFDLAS 365  
 397 IMTRGSSLMDEVNYGADIEGDMITTPRLPLIDMISNDPLIGYKLIANADAG-G 455  
 366 TLARELHDVDRLSAF-----FDLVQDDPPVVSQVKLIIEPMDVCEG 405  
 456 LYVGORPHMMVSEMNKRYDIYQFRTKGDGFAGFAECGSPHLYOAGGRKPMHSI 515  
 406 GYOVGNRP--GLWTEWNSKYDLYADYWRGEPALIGEFAISRLTSSDLEAGRRPSAI 463  
 516 NFVCAHOGFTLADLVYTKKYNLPNGENNROGHNHLSNCGEESGFARLSYKRLRKRM 575  
 464 NFVTAHOGFTLNDLVSYDKNKNEANGENNRGDESYNRSMNCVGEPTDDPLALARKM 523  
 524 RNMATLWVSGTSPMAIHAGDEIGTGYGNNNVYCODSELSYMDWSLVDNADLLAFARKA 583  
 636 TKRKE-----CEGLGLEDFTAKRLQW---HGHPGKPMSEN-SRVAASMKDE 682  
 584 TLLRKNHRRFRRRREFEGERINSDEVDIAWLTPSGRMTHEWDGREGDRCVAVFLNGE 643  
 683 -----RQGE-----IYVAFTSHLPAVVELP-ERAGRRPVEDVTGPR 719  
 644 AITAPDARGEKVVDLSFLICFNAHDHVEFVMPHDIYAQQTGELDTNDP 693

RESULT 2  
 GLGX\_HAEIN STANDARD; PRT: 659 AA.

AC P45178;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glycogen operon protein glgx (EC 3.2.1.-).  
 GN GLGX OR H11358.  
 OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., McInerney J.D.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/  
 CC CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U32815; AAC23005.1; -.  
 DR HSSP: H10342; 1BE2.  
 DR TIGR: H11358; -.  
 DR InterPro: IPR004193: Isoamylase\_N.  
 DR InterPro: IPR00461: Alpha.amylase.  
 DR Pfam: PF00128: alpha-amylase; 1.  
 DR Pfam: PF02922: Isoamylase\_N; 1.  
 KW Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.  
 FT ACT SITE 333 BY SIMILARITY.  
 FT ACT SITE 440 BY SIMILARITY.  
 SQ SEQUENCE 659 AA; 75290 MW; E2B71F610E8B1CAC CRC64;

Query Match 24.1%; Score 1014; DB 1: Length 659;  
 Best Local Similarity 33.9%; Pred. No. 1.4e-69;  
 Matches 236; Conservative 112; Mismatches 242; Indels 106; Gaps 21;

60 GMPAPLG-ATYALAGV--NFAVYSGGATAAALCLFTPEDLKADRYTEVPDPLMNRG 115  
 8 GMPAPLG-ATYALAGV--NFAVYSGGATAAALCLFTPEDLKADRYTEVPDPLMNRG 115  
 116 NMHNVFIEBELHNLGYRFEDGFAPHCCHYLDVSNVVDPAKAVISRGEGVPPARGN 175  
 59 NMHNLAVTVKGTETFAFRIHGEFA-----NPKRLIIDPYAKAA--NGKPLDSEES 109  
 176 CW-----POMAGMIPPLPYSTFPMESDPLRYPOKDLVIYEMHRCFTKIDSNVH 226  
 110 SMFLSDNDNNAHLAPRAVYVISEFPMENDTSNTPMAETIYELVYKFSQ---LNEKI 166  
 227 P---GTFIGAVS--KLDYLKELGVNCEILMPCHEFNELEYSTSSKMNFGYSTINF 280  
 167 PALRGTYTGLAHAPVNLAVLTKELGVAVVELLPVNFHINPRLDARGLQVNYMGVNLPMFA 226  
 281 PMRTYTSIGIKKCGRAINEFTFVREAKRGIEVILDVVFENFAGNENGPILSKGYD 340  
 227 VEPKYAA-----TNPLAEFTYKAFKAGIEVILDVVFENFAGNENGPILSKGYD 340  
 341 NTTYTTLARKEFYVNSGNTFNCNHPVYVQFIVDCRYVTEMHVDFGFFDLASIMTR 400  
 281 DQTYTWRNDQGRYINMTGCGNNLNTSSDYGKRKYVDCLATYWDQCHIDDFRDLATVLR 340

```

QY 401 GSSLMDPVNVYCAPLEGSMITTTGPTLVTPPLIDMISNDPLIGVKKLALAEAMDAGL-YQY 4355
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 341 DTPDQNS-----SAQLETDIKNEPSLQNIKKLLAEPMWDIGHXGYQ 380
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 460 GQFPHMNVMSWMNKRYRDIVROFTKGTDFGAFGAEECLCSGPHLYXQGGKPMHSMFYVC 519
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 361 GNFPSY--FAEDNDREFRDDLCREPLWMSGELGAFARERAGSSDLEFKKNDRLPHTTLFEIT 438
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 520 AHDGFTLADLVTYNKKYMLPGENNNRPGENHNLSPNCGEEGFARLS-----VKRLRR 573
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 AHDGFTLKDLYSYNDKHNETGNGEENRDRGRNNTSYNNGVEBSTLSLEPQKSAVENNRTE 438
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 574 QMRNFECYCLAWSQGVPMFEMGDEXGHTKGGNNNTYCHDSYVYFRRWDKBOYSELHRFCC 633
      ::::: ::::: ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 AOSGLMLSLANGTPMLLAGDEGNTQYQNNMAYCQDNETTWLKMNFNE--ELFELTK 556
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 634 LMTFRKCEBGLGLEDFFTAARLQWGHQPGKP---DMSNSRFVAFSMKDER----- 663
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 557 QTIALRKQIGSLNNDQWMSDENYQML-NIYGEPTVEDWQOQTKALQVLDNRMLLIN 615
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 684 ---QGEIYVANTSHLPAVVELLPBACGRMPVYDT 716
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 616 AKAEQME-----HLP-----NKKMPQDGT 636
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

RESULT	3			
GLGX_ECOLI		STANDARD:	PRT:	657 AA.
ID	GLGX_ECOLI			
AC	P15067: P76693:			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycogen operon protein glgx (EC 3.2.1.-).			
GN	GLGX OR GLYX OR B3431.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562;			
RP	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-K12 / MG1655;			
RC	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12."			
RL	Science 277:1453-1474(1997).			
RP	[2]			
RP	SEQUENCE OF 1-590 FROM N.A.			
RC	STRAIN-K12;			
RC	MEDLINE=89108020; PubMed=2975249;			
RA	Romeo T., Kumar A., Preiss J.;			
RT	"Analysis of the Escherichia coli glycogen gene cluster suggests that			
RT	catabolic enzymes are encoded among the biosynthetic genes".			
RL	Gene 70:363-376(1988).			
CC	-1- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/			
CC	CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO			
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	-1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A STOP.			
CC	CODON IN POSITION 507.			
CC	-----			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -			
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -			
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY			
CC	COMPROMISED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL			
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).			
CC	OR SEND AN EMAIL TO <a href="mailto:license@sib-sb.ch">license@sib-sb.ch</a> ).			
CC	-----			
DR	EMBL: J01616::AAA98735.1: ALT_SEQ.			

DR	EMBL:	U18997:	AAA58329.1;	-.
DR	EMBL:	AE000419:	AAC76456.1;	-.
DR	PIR:	JT0400:	BVECGX.	
DR	HSSP:	P10342:	1BF2.	
DR	ECogene:	EG10381:	glgX.	
DR	InterPro:	IPR000461:	Alpha_amyase.	
DR	InterPro:	IPR004193:	Isoamyase_N.	
DR	Pfam:	PF00128:	alpha-amyase; 1.	
DR	Pfam:	PF02922:	Isoamyase_N; 1.	
KW	Glycogen	biosynthetis;	Hydrolase;	Complete proteome.
FT	ACT_SITE	336	336	BY SIMILARITY.
FT	ACT_SITE	443	443	BY SIMILARITY.
FT	CONFLICT	288	288	Y -> Y (IN REF. 2).
SQ	SEQUENCE	657 AA:	73576 MW;	AA9EFC1F67FD0420 CRC64;

Query Match	23.8%	Score 1001	DB 1	Length 657
Best Local Similarity	34.4%	Pred. No. 1,3e+68		
Matches 235	Conservative 98	Mismatches 261	Indels 90	Gaps 18
QY	60	GMPADLGGATLALGAVNFVNSGSGATAAALCTLPEDLAKADRYTEEVPLDPLMRTGNVNH	119	
Db	7	GKPAFLGHHYDQGVNFTFLFSAHARVELCVF-----DANGGHRD-LGHSGDIGH	58	
QY	120	VFEISELHNMMLGYTFPDGTFAPHCCHYLIDVSNVVDPAKAV-----ISRGEYVPA	171	
Db	59	GFLPARBGLRGYGVHVGWQPAEBHRFPALKLIDPCARQIDGFEKDPRLHAGHNEDP	118	
QY	172	RGNMCPWQAGMIRLPYSNFDWEGDLPRLPYPOKDLVITYEMHRRGET-KHDSNNVEHPCF	230	
Db	119	YRDNA-AIAPKCVVVDHYDMEDDAPRPTMGSTITIEAHKGLYLHPELRIVELRGTY	176	
QY	231	--IGAVSKLDVLEKLGVCNCEILMPCHENNELYSTSSKMPNMGYSTINFSFPMTRYSG	288	
Db	177	KALGHPVMINYLKQIGITALLELPAVAFASBDRLOMGSLSNWGVNPAVMALHPAYA--	234	
QY	269	GIKNGRD-AINEFTFPRBAHKRCEIYLIDVFNHTABGNENGYLSKGYDNTYYML	347	
Db	235	---CSPELALDEFDAIKALHKAIEVILDLVLHNSAELDLGDLFSLRGIDNRSYYWI	290	
QY	348	APKGEFVNSGCGNFNCNHPVROFIYDCLRYVTTEMHVDGFRFDLASIMTGGSLMDP	407	
Db	291	REDGXYHNWWTGCGNTLNSHRAVVDYASACLRKYVETCHYDGFRRDLAAVGR-----	343	
QY	408	VNVYAPLEGDMITTTGTPLYTP-----PLIDMISNDPILGCVKLIAEAMD-AGGLXOYG	460	
Db	344	-----TPEFRQDAPLFLFAIONCPVLISGVKLIAEBWIDAPGGYVG	383	
QY	461	QEPHNWVSEWNGKRYDLYROFIKTDGPAFGAFACLCGSPHLYXOAGKRPKHSTINFCA	520	
Db	384	NFP--PLFAEMDNHFRDAAFRFMYLDLPLGAFAGRAAASDVFERNGLPFAAINTLVTA	441	
QY	521	HDGFTLADLVYNNKYULPENGNNNGENHNLNMYNMGEGEGEFARLSVYKLRKQMRNFV	580	
Db	442	HDGFTLRDVCYFNHKNHNEANGENBNDGNNNYNNHNGKEGGLGSLDIYBRRDSTHALLT	501	
QY	581	CLMVSQGVPRMFMDEGHTKGGNNNTYCHDSYVYFPMDKKEOYSELHRPCLMTKPRK	640	
Db	502	TLLLSQGFPRMLLAGBEGHSHOHGNNNAACODQLTWLDMSQAS--SGLTAFPAALIHNRK	559	
QY	641	ECEGGLGLEDFTAKRLQVHGHOPKGRPOVSENSRFAFASMKDEBROG-----ETIYV	689	
Db	560	R-----IPLAVENRKMEEBDCGVNRHL--NRYAQPLSTDEWONGPRQDQILSLDRFLI	609	
QY	690	AFNTSHLPAVVELPERAGRRMEPV	713	
Db	610	AINATLEVEITVLP--AG-EMHAI	630	
RESULT	4			
ISOA_FLASP		STANDARD:	PRT:	777 AA.
ID	ISOA_FLASP			
GC	032611;			

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Isoamylase precursor (EC 3.2.1.68).
GN IAM.
OS Flavobacterium sp.
OC Bacteria: CFB group: Flavobacteria; Flavobacteriaceae; Flavobacterium.
OX NCBI_TaxID=339;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97340931; PubMed=9197405;
RA Krohn B.M., Barry G.F., Kishore G.M.;
RT "An isoamylase with neutral pH optimum from a Flavobacterium species:
  cloning, characterization and expression of the iam gene.";
RL Mol. Genet. 254:469-478(1997).
CC -1- FUNCTION: HAS A HIGH RATE OF HYDROLYSIS FOR GLYCOGEN. DOES NOT
  CLEAVE PULULAN. HAS A PH OPTIMUM OF 6-7.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
  linkages in glycogen, amylopectin and their beta-limit dextrins.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U90120; AAB6356.1; -
DR HSSP: P10342; 1BP2.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR0004193; Isoamylase_N.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF02922; Isoamylase_N; 1.
DR KJ Hydrolyase; Glycosidase; Signal.
FT SIGNAL 1 32
FT CHAIN 33 777
FT ACT_SITE 410 410 BY SIMILARITY.
FT ACT_SITE 478 478 BY SIMILARITY.
FT ACT_SITE 533 533 BY SIMILARITY.
FT DISULFID 419 423 BY SIMILARITY.
SO SEQUENCE 777 AA; 84340 MW; 90548988BFD1445B CRC64;

Query Match 19.2%; Score 810; DB 1; Length 777;
Best Local Similarity 32.6%; Pred. No. 6e-54;
Matches 257; Conservative 96; Mismatches 271; Indels 164; Gaps 38;

OY 42 EYVAEDRYALGACRVL-----AGMFA-----PLGATRALAGVN--FAVYSGATA 85
DB 2 DPHAFQORSQGRRLALALALACALSPAHAIDAQOGLGARYDAQAQMLARFVYSSRAT 61
OY 86 AALCLF-TPED-----LKADRYTE-----EVLPLDPLMNTGNWNHFIQELHNLMTXG 132
DB 62 VEFLLYNPGSGQEVARLALSKDPATQWSLSLPTSTKNTYG-----ITGAVY---YG 112
OY 133 YRFDGTAPH-----CGHYLDVN-----VVVDPYAK-----AVISRGE 166
DB 113 YRANGPMPWIDAMNTKSGATGFSVDVNAQNRPNKLLIDPYAREISODPTATCAAGT 172
OY 167 -YGVPARGN-----CWPQAMGMIPLPYSTPDWEGDPLRYPKQDLVTEYHMLRGFTKHD 220
DB 173 IYANGAHRNKDGLGASKGIALAADATSV---GSKPTR-ALKDEVITYEVHVRGLTRND 227
OY 221 SS-NVEHPGFTIGAVSKLDYLAKEGVNCIELMPCHEF-----NELEYSSSSKMNFWGIST 275
DB 228 DSVPAARGYKGAARNAALAAAGVAVVEFLPVQETQNDONDVD--PSTAGDNTWGYMT 286
OY 276 INFSPMTRY---TSGIKKCGDAINEFKTFVREAHKRGIEVLVDVFNHTAGENG 331
DB 287 LNYAPDRRYAYDKSAGG-----PTREKMAKMAKAFHDAGIKYIDVYVYHNTGSG---G 336

```

```

OY 332 P-----ILSFKGVNTTYNMLAPKGEF-YNYSGCCGTFENCNHPVYRQFIYDCLR 379
DB 337 FWSGTDGLSVYNLSLFRGLDNPAYSLSSDYKYPVDNTGVGVGNTRHPIAONLIVDSLA 396
OY 380 YWVTEHVDGFRFDLASITMRGSSSLMDPVNNGVAPIEGDMITTGPLVTPPLIDMISNDP 439
DB 397 YWRDALGYDGRFPLDASLT--GNSQHCFCFNDKDKDGSNALNR-----IYALFPRPP 446
OY 440 ILG--GVRLIAEAMDAG-LYOVGQFPHMNVWSENGKRYDIR--QFIKTGDFR-GGF 493
DB 447 AAGGAGADLLIAPNALIGNSYQVGGEPAG--WAEKNGLYRDALRKONKLGVEYVTPGL 504
OY 494 AELCGSHLYQAGGRKRWHSINPYCANDGFTLADLVYNNKYN---LPNGENNDGENH 550
DB 505 ATRFAGSNDLYGDGDRKPMWSINFEVADGFTLNDLYVNDKONQNPWYGFSD--GGEHD 563
OY 551 NLSNMGEGEGERPARLSVRLRKROWRNFVCLMVSQGVPMFVMEGDEYHTKGNNTYCH 610
DB 564 NLSNMG-----GIYAEOKKARTGLALLMLASAGVPMITGDEBALRTQGNNTYVL 615
OY 611 DSYVYFMRDKREOYSELHRECCMTKRECEGLDEDPPTA-----RRLQMHQ 662
DB 616 DSAAMWLXMSALEADHETTKRLIAFRKAHPALRPANFYASDPTNGVMQLRMF--- 672
OY 663 PKPDMSE-----NSRFVAFSMKDERGE---IYVANTSHLPAVVELP-ERAGR 708
DB 673 --KPDGAADSAFYNGADNHALMWRIDSEFGDSASAIYVAVNGMSGVAFRLPMPGTCK 730
OY 709 RMEPVVD 716
DB 731 QMYRTDT 738

RESULT 5
ISOA_PSESP ID ISOA_PSESP STANDARD. PRT: 776 AA.
AC P26501;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Isoamylase precursor (EC 3.2.1.68).
GN IAM.
OS Pseudomonas sp. (strain SMP1).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.
RX MEDLINE=89381677; PubMed=2778432;
RA Tognoni A., Carrera P., Galli G., Lucchese G., Camerini B.,
  Grandi G.;
RT "Cloning and nucleotide sequence of the isoamylase gene from a strain
  of Pseudomonas sp.";
RL J. Gen. Microbiol. 135:37-45(1989).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
  linkages in glycogen, amylopectin and their beta-limit dextrins.
CC -1- INDUCTION: BY MALTOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: M25247; AAA2585.1; -
DR EMBL: A10909; CAA00929.1; -
DR PIR: A37035; A37035.
DR HSSP: P10342; 1BP2.
DR InterPro: IPR000461; Alpha_amylase.

```

```

DR InterPro: IPR004193; Isoamylase_N.
DR Pfam: PF00128; alpha-amylase_1.
DR Pfam: PF02922; Isoamylase_N_1.
DR Hydrolase: Glycosidase; Signal.
FT SIGNAL 1 26
FT CHAIN 27 76
FT ACT_SITE 401 401 ISOMYLASE.
FT ACT_SITE 481 481 BY SIMILARITY.
FT ACT_SITE 536 536 BY SIMILARITY.
FT ACT_SITE 410 422 BY SIMILARITY.
FT DISULFID 546 616 BY SIMILARITY.
FT DISULFID 738 766 BY SIMILARITY.
SQ SEQUENCE 776 AA; 83656 MW; A5E4C02EF026A3A4 CRC64;

Query Match 19.1%; Score 803.5; DB 1; Length 776;
Best Local Similarity 32.0%; Pred. No. 1,9e-53;
Matches 242; Conservative 99; Mismatches 289; Indels 127; Gaps 31;

OY 55 CAVLAGMMPA-----PLGNT--ALAGVVFVYSSGATAALCLFTPEDLKADRV 102
DB 13 CAVLAGVPPAPMAHAISMSLSGASDAQOANITFRVYSSQATRIYLYLS-----AGYGV 67
OY 103 EEPVLDPLNNRTGNMWHVF-----EGELHMLCYRFDGTFAPHC----- 143
DB 68 QESATYTLSPAGSGVAIVYVPSISAKAGITGAVYGYRAMGFNWPPLASWKGSGQAGFV 127
OY 144 -----GHYLDVSNVVDPAKAVISRGEYGPARGNNCW-----POMAG 182
DB 128 SDVDANGDRFNPKLLDPAQEV--SQDPLNPNONGNVFASGASYRTTDSGIAPRGVY 166
OY 183 MLEPLSTEDWEGDLELRIPQKDLVYEMHLRPFTHKDS--NVEHGETFGAVSKLDYLK 241
DB 187 LVPSTOST---GTRKPTR-AOKDVIYEVRHGFTEEDTSPAPQRYRTYGCAGKASYLA 241
OY 242 ELGVNCTELMPCH-----ENELEYSTSSSKMNFSGYSTIFSPMRY-----TSGGKNC 293
DB 242 SLGVAVERLEPVOETONANDV--VPNSDANONWGTMTENTFSPDRRYAVNKAAG---- 296
OY 294 GRDAINEFTFYREAHKRGILEVLDVVFNHTAGN-----ENGPILSEFGVNTTYM 346
DB 297 ---PTAEFGAVOAFHNAGIKVYNDVYVNTAGGTWTSDDPTATITYSRGIDNTTYE 353
OY 347 LAPKGE-FVNYSGCGTFCNHPVROFIYDCLRYVTEHVDGFRDLASIMTGGSLW 405
DB 354 LITSGNOYFYDNTGIGANFTYNTVAONLIVDSLAWANTMGVDFREDLASVL--GNSCL 411
OY 406 DPNVNYGAP--IEG--DMITGTPLVTPPLIDMISNDPLG--GVKLIARAMAGG-LYQ 458
DB 412 NCATYASAPNCPNGCYNFDAADSNAINRLREFTVRPAAGSGGLDFAEPPAIGGNSYQ 471
OY 459 VQGFPHMNVSEWNGKYYRDIYRQF---IKGTDFAGGFAECLCGSPHYOAGRKPPWHSI 515
DB 472 LCGFPGG--WSEWNGLFRLDSLROAONELGSMITVYQDANDFGSSNLSLFGSSGRSPWNSI 529
OY 516 NVCVAHDSGTLDLVYNNKYN---LPNGENNDGEMHNSMNGCGEGEFAKLSVRLRK 572
DB 530 NIDVDHGWTLTDVYSCNANNSQAMPYGPSD--GGSTJNSW---DQMGAGCAVADOR 585
OY 573 ROMRNFVCLMYSOGVPMFYMGDEYHGTGKGNNTYCHDSYVYFRMDKKEQYSELHRC 632
DB 586 RAARFGMAFEMLSAGTFLMGGGDEYLRITQCNNAVAINLDDSSANWLYYSWTTDOSNFTTFA 645
OY 633 CLMTFRKECEGLGLEDFTAKRLQWHGHPKRPD---WSENSRF-VAFSMDKDEGGE-- 686
DB 646 QRLIFRKAHPALRPSSWYSGSOLTWYQPSGAVADSNTYMNNTSNVAIAVAINPISLGDSDN 705
OY 687 -IYVAFN-----TSHLPVAVELPERAGRRMEVVDVT 716
DB 706 STYVAVYNGWSSVTFLLPA-----PPSGTQWYRVTDI 737

RESULT 6

```

```

ID ISOA_PSEAY STANDARD; PRT; 776 AA.
AC P10342;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Isoamylase precursor (EC 3.2.1.68).
GN IAM.
OS Pseudomonas amylocleramosa.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=32043;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=SB-15;
RX MEDLINE=88243808; PubMed=3379068;
RA Amemura A., Chakraborty R., Fujita M., Nouni T., Futai M.;
RT "Cloning and nucleotide sequence of the Isoamylase gene from
RL Pseudomonas amylocleramosa SB-15."
RL J. Biol. Chem. 263:9271-9275(1988).
RN 2;
RP SEQUENCE FROM N.A.
RC STRAIN=JD210;
RX MEDLINE=91064385; PubMed=2248978;
RA Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H.;
RT "Nucleotide sequence and expression of the isoamylase gene from an
RL Isoamylase-hyperproducing mutant, Pseudomonas amylocleramosa JD210."
RL Biochim. Biophys. Acta 1087:309-315(1990).
RN 3;
RP SEQUENCE OF 744-776 FROM N.A.
RC STRAIN=SB-15;
RX MEDLINE=89327147; PubMed=2753857;
RA Amemura A., Fujita M., Futai M.;
RT "Transcription of the isoamylase gene (iam) in Pseudomonas
RL amylocleramosa SB-15."
RL J. Bacteriol. 171:4320-4325(1989).
RN 4;
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=98387895; PubMed=9719642;
RA Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;
RT "Three-dimensional structure of Pseudomonas Isoamylase at 2.2-A
RL resolution."
RL J. Mol. Biol. 281:885-897(1998).
CC 1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
CC linkages in glycogen, amylopectin and their beta-limits dextrans.
CC 1- INDUCTION: BY MALTOSE.
CC 1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03871; AAA25854.1; -
DR EMBL: X13378; CAA31754.1; -
DR PIR: A28109; A28109.
DR PDB: 1BF2; 12-AUG-98.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR004193; Isoamylase_N.
DR Pfam: PF00128; alpha-amylase_1.
DR Pfam: PF02922; Isoamylase_N_1.
KW Hydrolase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 76
FT ACT_SITE 401 401 ISOMYLASE.
FT ACT_SITE 481 481 BY SIMILARITY.
FT ACT_SITE 536 536 BY SIMILARITY.
FT DISULFID 410 422 BY SIMILARITY.
FT DISULFID 546 616 BY SIMILARITY.
FT DISULFID 738 766 BY SIMILARITY.

```





CC	-I- MISCELLANEOUS: THIS ENZYME IS EXTREMELY THERMOSTABLE.
CC	-I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).
CC	
DR	EMBL; D10487; BAA01368.1; .
DR	PIR; A41707; A41707.
DR	HSSP; P21332; 1UOK.
DR	InterPro; IPR000461; Alpha_amylase.
DR	Pfam; PF00128; alpha-amylase; 1.
KW	Hydrolase; Glycosidase.
FT	ACT_SITE 199 199 BY SIMILARITY.
FT	ACT_SITE 256 256 BY SIMILARITY.
FT	ACT_SITE 330 330 BY SIMILARITY.
SQ	SEQUENCE 562 AA; 66505 MW; 925EB5924726D42A CRC64;
Query Match	7.4%; Score 310.5; DB 1: Length 562:
Best Local Similarity	21.6%; Pred No. 4.7e-16;
Matches 144;	Conservative 98; Mismatches 168; Indels 257; Gaps 32.
OY	204 KDLVYEMLRGFTKHDSSNVHPGTIGAVSKLDYLKELGVNCIELMCHFEFNELEYST 263
Db	7 KAAYVYQIYPRSFY---DSNGDIGIDIRIIAKLDLYKELGADVVMLSVYRK----- 55
OY	.264 SSSKNMFNGYSITNFPSPRTTRTSGCINKCGDAINER-----KTVEAHKRGIEVIL 317
Db	56 SPNDNN--GYDISDY-----RDIDEFGTMADMTMLEMKHKGIKLVW 97
OY	318 DVVFNHTEAGNENGR--ILTFKVGN--TTYWMLNP----- 349
Db	98 DLVYVHTSDSEH---FWLFESKRSKSNPRDIYIMRPKCKGKEPPNMVESVSCSAEYDEM 154
OY	350 KGEFY--NVSGGNTFNCHNPPYROFIYDCLERYVTENMHVDGFRRDLASINTRGSSLMP 407
Db	155 TGEYTLHLFSKKQPOLDLNEMNPKVRREYEMKFMFL-DKGVDFRMDIVYMISKVPELD- 212
OY	408 VAWYGAPLEG-----DMITTG-TPLVTPPLDMMT 435
Db	213 ---GEPSGKKYVASGSRYNMGPRVHFLEOMNREVLSKYDIMTVGETPGVTPK-EGIL 267
OY	436 SNDPLTGLVKLIAEWDAGLYGVGFOPFHMMNWMSMNCKYBDIVYRQFIKGDGFAGRAE 495
Db	268 YTDP-----SRRELNMVFOFHEIMDDSGSGKW-DI----- 297
OY	496 CLCGSPHLYOAGRKPMHSINFVACHADFETLADIATYTKKYLNPGENNRDGENHNLSMN 555
Db	298 -----RPW-----SLADL-----KTTMKRWQKELESG---GMN 322
OY	556 C-----GEEGEFARLSVLRRLKRROMRNFVCLAMSOGVPMFYMGDEHGHT 600
Db	323 SLYLNNHDOPAVRSRFEGDKRY-----RVESAKMIAFLHMM-QGPRIYIGBELGMT 374
OY	601 KGNNNTYCHDSYNYFPWKDKEDQYSELHRCCLMTFRKRCCEGLGLEDFPTARKLOHG 660
Db	375 N-----VAFPSIEDYRDE---TLNMRKEVVEEGDEPDQEVMEKRIYYKG 415
OY	661 HQ---PEKPMSENSRFVAFSMKDERGEGEIVYAFNTSH---LPAYVELPERAGRMRP 712
Db	416 RDNARTPMQWMDSENAGTA-----GTWMIVYNPNRYKRIANKALAEEDNSVFHYKK 467
OY	713 VYVDGKP-----APYDFLTDDL-----DRALTTHOPS-----HFLYS 745
Db	468 LIQRKHODIIVGYTDLILEDDEVIYRYTTRLIGNEQLIIVTNSESEKTPVFRPLDHIIYK 527
OY	746 NLYPMLS 752



SQ SEQUENCE 596 AA; 65262 MW; 73EA80AE0534DDCD CRC64;

## Query Match

Best Local Similarity 27.0%; Score 283.5; DB 1; Length 596;  
Matches 138; Conservative 53; Mismatches 147; Indels 173; Gaps 31;

```
QY 131 YGAFEDGFAFHGCHYLDVSNVVPYAKAVSRGEYCPA-----RNNKCPQWAGMT 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 YGILLDGEIP-----LPPRPR-----ROPEGHALSRFPDPAHFM-QDAG-- 106
QY 185 PLPYSTFMEGDLPLRYPOKDL--VIYEMHLRGFTKHDSSNVEHGPTEIGAVSKLDYK 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 107 -----WQG-----RELQGSVITYELHIGTFTPE-----GITDAAGKLDYLA 142
QY 242 ELGVNCTELMPCHEFELEYSTSSSKMNFWGYSTINFSPMRYTSGIKNGRDAINEF 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 143 GLGIDFIELLPVNAFN-----GTHNMGYDVQWFAVHEGY--GCPA-----AY 183
QY 302 KTFYREAKRGIEVILDVVENHTA-EGN---ENGPILSKGYDNTTYMLAKGEFYNS 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 QREYDAAMAGLGYLDVYVNHLLGPSGNYLPRYGYLK-HGEGNTW-----GDSVND 235
QY 358 GCGNTFNCNHPVROFIYDCLRYVWTEMHVDFEEDLASIMTSGSSIMDPVNVYCAPIEG 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 GPGS---DH--VKQYILDNVAAMLRDVRDGLRDAVAHAL-KDERAVHILEEFGA---- 284
QY 418 DMITGTPPLVTPPLIDMISNDPILGG--VKLIAEA-----WDAGLTYVQGF-- 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285 -----LADALASSE--GGRPLTLIAESDLNNRLLYRDVNVGYLAQMSMD 328
QY 464 -----HNVSEWNGKYKD-----IVROFIKTDGFGAGFECLSGSPHYQAGGR 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 FHAAVHVAVSETTGYSDFSLGALAKVLR-----DGF---FHD---GSYSFPR--GR 374
QY 510 KPMHSINVCAGHDGFTLADLYTNKKYNLPNGENNRDGENHNLISNCGEGEFALSYKR 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 CHGPRINFSAVHP-----AALVYCSNHDQIGNRATGDRLSQSLPY-----GSLAALAVLT 425
QY 570 LRRKROMRNFVCLANVSQGVPMFYMGDEYGH 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 L-----TGPRTPMLFMGEYGAJ 443

RESULT 11
TREZ_ARTRM STANDARD: PRT: 575 AA.
ID TREZ_ARTRM
AC 09AJN6;
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DE MaltO-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-((1->4)-alpha-D-glucano)trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Archaeobacter ramosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthropacter.
OX NCBI_TaxID=1672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S34;
RA Yamamoto T., Murata K., Watanabe H., Yamashita H., Kubota M.,
RA Fukuda S., Kurimoto M.;
RT "trehalose producing operon trezA from Arthropacter ramosus S34."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-((1->4)-alpha-D-glucanosyl)(n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -I- PATHWAY: trehalose biosynthesis.
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC [1]
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB045141; BAB40766.1; -  
DR InterPro; IPR000461; Alpha\_1\_amilase.  
DR Pfam; PF00128; alpha-amyase; 1.  
KW Hydrolase; Glycosidase.  
FT ACT\_SITE 250 250  
FT ACCT\_SITE 250 250  
SQ SEQUENCE 575 AA; 63079 MW; 284A3F20207E228B CRC64;

## Query Match

Best Local Similarity 25.9%; Score 278; DB 1; Length 575;  
Matches 136; Conservative 56; Mismatches 178; Indels 156; Gaps 29;

```
QY 104 EYPLDPLMRNTGNVHV---IEGELHNLMLGYRFGTFAPHGCHYLDVSNVVPYAKA 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 ELPL-----TRDNGMVALDQPPMDGPDLYDGYLDGK-----GFPADP 64
QY 161 VISRGEYGVARGNMCWPMAGMIPLYSTPD--WEGDLPLRYPOKDL--VIYEMHLRG 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 RSLRQPRGVHEIGRE-----FDPARYAMGDGMRG-----RDLGAVITYELHVG 109
QY 216 FTKHDSNVEHGTPLGAVSKLDYKELQVNCLELMPCHFEFNELEYSTSSSKMNFYST 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 FTPE-----GTLDSALRRLDHLVRLGDVAVELLPVNAFN-----GTHGNGYDG 152
QY 276 INFSPMRYTSGIKNGCRDAINEKTFVREAKRGIEVILDVVENHTA-EGN---ENG 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 VLMYAVHEPY--GGRP-----AYGRFYDACHANGLAVDVNVNHLLGPSGNYLPRDG 202
QY 332 PILSKGYDNTTYMLAKGEFYNSGCGNTFNCNHPV---VROFIYDCLRYVWTEMHVD 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 203 PLYG--SGAANT-----MGDALNLDPRLSDEVARYIIDNAVYVLRDMHMD 245
QY 389 GRPPLASIM--TRGSSIMDPVNVYCAPIEGDMITGTPPLIDMISNDPIL----- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 GLRLDVHVALRDARALHLEELAAVYDLAELGEL--GRPLTLIAESDL--NPKLLRSRA 300
QY 442 -GGVRLIAEAMPDAGLYGVGQEPHNVSEWNGKYKD-----IVROFIK--TDGFAAG 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 AHGYGLDAG-WD-----DVHNAVHNAVGTGYADFGCLAKLVKVFQGRFHDGTSS 355
QY 493 FAECLGSPHYLYQAGGRKPMHSINVCAGHDGFTLADLYTNKKYNLPNGENNRDGENHNL 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 FRRHRHGRPLDIPFRR---LVAFQDHQV-----GNR-AVGDMR 393
QY 553 SNVCGEGEFALSYKRLLKRRKROMRNFVCLANVSQGVPMFYMGDEYG 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 SAQVG-EGSLAAA-----ALVLLGPRTPMLFMGEEMG 425

RESULT 12
TREZ_BREHE STANDARD: PRT: 589 AA.
ID TREZ_BREHE
AC 052520;
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DE MaltO-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-((1->4)-alpha-D-glucano)trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Brevibacterium helvolum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Brevibacteriaceae; Brevibacterium.
OX NCBI_TaxID=1704;
RN [1]
```

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 11822:  
 RA Kwon T., Kim C.H., Choi Y.D.:  
 RT "Maltotriose synthase/trehalose hydrolase encode proteins for  
 RT trehalose production in *Brevibacterium helvolum*."  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic  
 CC linkage in 4-alpha-D-(1->4)-alpha-D-glucanosyl(n) trehalose to  
 CC yield trehalose and alpha-(1->4)-D-glucan.  
 CC -1- PATHWAY: trehalose biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----  
 CC EMBL: AF039919; AAB95369.1;  
 DR InterPro: IPR000461; Alpha-amylase.  
 DR Pfam: PF00128; alpha-amylase; 1.  
 KM Hydrolyase; Glycosidase.  
 FT ACT\_SITE 258  
 FT SEQUENCE 589 AA; 64217 MW; C282314B6E9BA029 CRC64;

Query Match 6.6%; Score 277.5; DB 1; Length 589;  
 Best Local Similarity 26.9%; Pred. No. 1.6e-13;  
 Matches 132; Conservative 61; Mismatches 161; Indels 137; Gaps 29;

OY 131 YGREDGTGAPHCGLDVSNVVDPAKAVISRGEYGPARGNCMPQAGMIPLEYST 190  
 DB YGRLDGNTP-----VPEPRSR-----IPAGYN--HSRTNPPRYM 95  
 OY 191 PD--WEGDPLRLRYPOKDL--VYEMHLRGFTKHDSNNVHETFGAVSKLDLYLKE 245  
 DB 96 QSRMNG-----KELOGTLYLQHV-GTSTPD-----GTLDAAGEKSLYLDVGI 139  
 OY 246 NCIELMPCHFEFNELESTSSKMFNGYSTINFSPMTRTSGGIRKNGRDALNEKFEV 305  
 DB 140 DEIELLPVNGFN-----GTHNMGYDQVQWYTHNEG--GGPA-----ATQREY 180  
 OY 306 REAHKRGIEVLIDVFNHTA-EGN---ENGPISEKGVNDTTYMLAPKGEFYNSGCGN 361  
 DB 181 DAHAAGGLGVIDVYVNHILGRLGNYPKLGPNLK--OGDANTL-----GDSVNLDDAGS 232  
 OY 362 TENCNHPVROFLVDCLRKWTYEMHVDGFRDLASIM--TRGSSLMDPVNVGAPTEGDM 419  
 DB 233 -----DYEREYILDNAALWVGVDYHVDGVDVAVHVRBRAVHILLEDLGLDADISGE- 285  
 OY 420 ITTGTPPLVPPRLIDMISNDPILGVLKLAEMADAGLYQVQPP-----HNNVSEWN 472  
 DB 286 --TGLPKTL--IAESDFNP-----RLIYPR-DVNGCYGLAGQWSDDEHTAVHVSSETT 335  
 OY 473 GKYRD-----IVROPIKGTDFGAGFAECLCGSPHYQAGCKRPMWSINFCADHGTLLA 527  
 DB 336 GYSDPESLAVLAKVK--DGFLHD-----GSYSFR--GRHHGRIPNLSLANP---A 381  
 OY 528 DIVYTNKKKNNLNGENNRGENNHSWNGCEGEFARLSKRLRKQMRNFECWVWSOG 587  
 DB 382 ALVVCQNNNDQJGNRATGRLSOSLSY-----GOLVAVALTL-----TSPF 423  
 OY 588 VPMFYMGDEYG 598  
 DB 424 TPMLFNGEYEG 434

RESULT 13  
 TREZ\_MYCTU STANDARD; PRT; 580 AA.  
 ID TREZ\_MYCTU

AC 010769;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Maltotriose synthase/trehalose hydrolase (EC 3.2.1.141) (MTHase) (4-  
 DE alpha-D-(1->4)-alpha-D-glucano]trehalose trehalohydrolase)  
 DE (Maltotriose) trehalose trehalohydrolase).  
 GN TREZ OR RV1562C OR MT613 OR MTCY48.03.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of *Mycobacterium tuberculosis* from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Ormazabal L.A., Esmailova M.D., Salzberg S.L.,  
 RA Delcher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic  
 CC linkage in 4-alpha-D-(1->4)-alpha-D-glucanosyl(n) trehalose to  
 CC yield trehalose and alpha-(1->4)-D-glucan.  
 CC -1- PATHWAY: trehalose biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----  
 CC EMBL: Z74020; CA98329.1;  
 DR EMBL: AE007027; AAK45880.1; ALT\_INIT.  
 DR TIGR: MT613;  
 DR TubercuList; RV1562C;  
 DR InterPro: IPR000461; Alpha-amylase.  
 DR InterPro: IPR004193; Isoamylase\_N.  
 DR Pfam: PF00128; alpha-amylase; 1.  
 DR Pfam: PF02922; Isoamylase\_N; 1.  
 KM Hydrolyase; Glycosidase; Complete proteome.  
 FT ACT\_SITE 247  
 FT SEQUENCE 580 AA; 64076 MW; 506190468F48B62 CRC64;

Query Match 6.5%; Score 272; DB 1; Length 580;  
 Best Local Similarity 21.3%; Pred. No. 4.2e-13;  
 Matches 144; Conservative 83; Mismatches 220; Indels 230; Gaps 36;

OY 118 WHVFEGELHNLGYREDGTGAPHCGLDVSNVVDPAKAVISRGEYGPARGNCNM 177  
 DB 33 WHTVAAPA-DARYGYLD-----DDPTVLPDPRS-----ARPGVHARSQRWE 76

```

OY 178 POMAGMILPSTFEDMEGDLPLRYPOKDLVYEMHLRGFTKHDSSNVEHPTGIVASKL 237
DB 77 P-PGOGFARITDGMPC-----RSVEGAVIYEHITGFTT-----AGTEPAALTEL 121
OY 238 DYKLEIGVNCIELMPCHEFELESTSSKMNFWGYSTINEFSPMTFRTYKINCGDA 297
DB 122 DYLVLDGIDVELMPCVNSF-----AGTRGNGYDGLMYSVHERY--GGPRGLV-- 168
OY 298 INEFKTFVREAHKRGIEVILIDVFNHTA--EGN--ENGPILSFKGVNTTYMLAPKGEF 353
DB 169 -----FIDACHARRGLVLIDAVFNHLGPGNLYPRFGPYLS--SASN-----PMGDG 213
OY 354 YNSGGGNTFNCNHPVROPTIVCLRWVYEMHVDGFRFLASLMTGSSLMDEVNYGA 413
DB 214 INIAG-----ADSEVRRHIIIDCALRMRFHADGLRLDAVHALVDTTAVHVELEANA 267
OY 414 P-IEGDMITGTPLVTPPLIDMISNDPILGKVLIAEMDAGGLYGVGPRHNNVSEW 471
DB 268 TRWLSGOL--GRPLSL--IAETDRNDP-----RLITRPSHG--YGTAAQW 307
OY 472 NGKYRDIVROFIKTDGFGAGFAECLOGSPHLXOAGGRKPMWSINFYCAHDEFTLADLV 531
DB 308 N-----DIIHAI-----HTAVSGERGYYA-----DFGSLATLA- 337
OY 532 YNKKYVLPNG-----ENNBDGE-----NHNLSMNCGESEGFARLSVKRLRK 572
DB 338 -----YLRNGYFHAAGTYSFRRRHRGALDTSALPATRLAYTCTHD-----QVGNRALGD 389
OY 573 RQMR-----NFVCLMWSOGVP--MEYMGDEGHTKGNNNTYC-----HDSYVNY 616
DB 390 RPSQYITLGGOLAIKALITLSPYTAALFKEBMC--ASSPFOFCSHPEPELAHSTVAG- 446
OY 617 FRWDKKEQYSELHRRFCCMLTKFRKECEGLDEFPYAK-----RLQWHGHPGRK-- 665
DB 447 -----RKEEFAR-----HGWADDDIPPODQTFQRCKLNNMAEASGSHAR 487
OY 666 -----PDSENSRFPYAFSKDERGCELIYVAFTSHLPA 698
DB 488 LHRFYRDLIALRHNEADLADPWLHMLMYDYDEQRRVYM-----RQOLMIACTMGAEP 542
OY 699 VVELPERAGRWE--PVV 714
DB 543 CVPVSGELVLAWEPII 559

RESULT 14
PULA_KLEPN STANDARD: PRT: 1090 AA.
AC P07206;
DB 01-APR-1988 (Rel. 07, Created)
DB 01-APR-1990 (Rel. 14, Last sequence update)
DB 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
  glucosidase) (Pullulan 6-glucanohydrolase).
CN PULA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
  Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UNF 5023;
RX MEDLINE=90205629; PubMed=2181242;
RA Kornacker M.G., Pugsley A.P.;
RT "Molecular characterization of pula and its product, pullulanase, a
  secreted enzyme of Klebsiella pneumoniae UNF5023.";
RL Mol. Microbiol. 4:73-85(1990).
RN [2]
RP SEQUENCE OF 1-62 FROM N.A.
RX MEDLINE=86033621; PubMed=3902792;
RA Chapon C., Raibaud O.;
RT "Structure of two divergent promoters located in front of the gene

```

```

RT encoding pullulanase in Klebsiella pneumoniae and positively
RT regulated by the malt product.";
RL J. Bacteriol. 164:639-645(1985).
RN [3]
RP SEQUENCE OF 944-1090 FROM N.A.
RX MEDLINE=89291709; PubMed=2661532;
RA D'Enfert C., Pugsley A.P.;
RT "Klebsiella pneumoniae pULs gene encodes an outer membrane
  lipoprotein required for pullulanase secretion.";
RL J. Bacteriol. 171:3673-3679(1989).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
  linkages in pullulan and in amylopectin and glycogen, and the
  alpha and beta-limit dextrans of amylopectin and glycogen.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
  (Probable).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb.ch/announce/
  or send an email to license@isb.ch).
CC -----
DR EMBL: X52181; CA36431.1; -
DR EMBL: M12503; AAA25087.2; -
DR EMBL: M29097; AAA61976.1; -
DR PIR: A25025; A25025.
DR PIR: A32880; A32880.
DR PIR: S11823; S11823.
DR InterPro: IPR000461; Alpha amylase.
DR InterPro: IPR004193; Isoamylase_N.
DR Pfam: PF00128; alpha-amylose; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hydrolyase; Glycosidase; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 1090
FT LIPID 20 20
FT ACT_SITE 684 684
FT ACT_SITE 713 713
FT ACT_SITE 841 841
FT CONFLICT 6 6
FT CONFLICT 10 10
FT CONFLICT 15 15
FT CONFLICT 23 23
FT CONFLICT 31 31
FT CONFLICT 34 34
FT CONFLICT 36 36
FT CONFLICT 55 58
SQ SEQUENCE 1090 AA; 118098 MW; 240AE7DFB3F1BD6 CRC64;

Query Match 6.38; Score 265; DB 1; Length 1090;
Best Local Similarity 19.78; Pred. No. 3.4e-12;
Matches 151; Conservative 89; Mismatches 244; Indels 284; Gaps 35;

OY 21 GEVCA-----AVEAATKVEDGEDEDPVADRYALGACRVLAGMPAPGATAGAGV 74
DB 266 GETVAIAAEDGILISATVOGTAGVLDAVAEAAALS-----YGAQLADGCV 313
OY 75 NEANYSGATRAALCLTPEDLKADRYTEVPLDPLNRRGNVHVFTIEGELHNNLYGR 134
DB 314 TFRWMAPTAOQVDVYVSAD---KKVIGSHFM--TRDASGASMSWOGSGDLKCAFYRYA 367
OY 135 FDGTFAPHCYHLDVSVWVVDPYAKAVISREGEVVPARGNNCWPMGMPIPLSTFPMW 194
DB 368 MT-VYHPOSKRKEVQYE--VTDPIAHSLSTNSYS-----QVVDLNDLSALKPDGMD 414
OY 195 GDPLRLRYPRK-----DLVITYEMHLRGFTKHDSS-NVEHPTFTI-----GAVSKLDLYLKE 243

```

Db	415	-NLTPHAQKTKADLAKMTTHESIINDLSAMDVTVAELRGKLTALTADDSMMVOHKLTL	473
Qy	244	--GVNCIEIMP-----CHEF-----	256
Db	474	SASGVTHVELLPVDFLATVATNEFSDKVDIDQPPSRICEVNSAVKSEFAGYCDGSJYEE	533
Qy	257	--NLELESTSSSKNF-----WGISTINFEFSPMTRRYISGCKKCGDAINE	300
Db	534	VLNOLKSDSDQDNPDVOALNTLVAQTDSDYNGWGDPPHYVPEGSVATDPE--GTTRIKE	590
Qy	301	EKTVPRAHRK--GIEVILDVFNHPTAEGNENPILSPFGVDNNT-----YMLAPK	350
Db	591	FRTHI-QAIKODLCMYIMDVYINHT--NAAP-----TDRTSYLDKIVEMWYQ----	636
Qy	351	GEFYNSGCGTENC-----NHPVVRQFIVDCLRYWTEMHVYDGFRLDASIMTRGS--	402
Db	637	-RLNETTGSVESATCCSDSAPRHRMFAKLIADSLAVWTDYKIDGRFPDLMGYHPQAOL	695
Qy	403	SLMPRVANVYAPRIEGDMITTTGTLVYLPPLIDMISNDPILGKYLLEAMDAGGLYOVGF	462
Db	696	SAWRIKALNPDI-----YFEGEGWDS-----NQS	720
Qy	463	PHMWVSEMN-----GKYRDIYRQFIKG-----T	486
Db	721	DRELIASQINLKGTGIGTFESDRCLDSYRGGGGPPDSGDALRQNGICSGAGVLPNELASLS	780
Qy	487	D-----GFAGGFAECLC-----GSPHLYOAGGRKPWMSINEV	518
Db	781	DDQVRHLADLRLCGMAONLADPFVIMDKDGAKKGSEIDYNGAGGTAA--DPEVYANV	837
Qy	519	CAHDFGLTADLVYTNKRYNLPNGEENNDEGNHNLNMCNCEGEGFARLSVKRLRKQRNF	578
Db	838	SKHNDQFLMDIMSY-----KASQEAFLA-----TRVRQAVS	869
Qy	579	FVCLAMVQGVPMFYMGEDEYGTGKGNNTYCHDSYVNYFRMDKKEYS	626
Db	870	LATVMLQGLIAFDQGGSELLRSK-----SPTROSY-DSGDGNFRVDYS	911
RESULT	15		
NEPU_THEVU			
ID	NEPU_THEVU	STANDARD;	PRT; 585 AA.
AC	Q08751;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, last sequence update)		
DT	15-JUL-1999 (Rel. 38, last annotation update)		
DE	Neupululanase (EC 3.2.1.135) (Alpha-amylase II).		
DN	TYAIL.		
OS	Thermoactinomyces vulgaris.		
OC	Bacteria; Filimicroles; Bacillus/Clostridium group; Thermoactinomyces.		
OX	NCBI_TaxID=2026;		
RN	11		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-R-47;		
RC	MEDLINE=93222535; PubMed=7763540;		
RA	Tonozuka T., Ohtsuka M., Mogi S.-I., Sakai H., Ohta T., Sakano Y.;		
RA	"A neopululanase-type alpha-amylase gene from Thermoactinomyces		
RT	vulgaris R-47."		
RL	Biosci. Biotechnol. Biochem. 57:395-401(1993).		
RN	12		
RN	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).		
RP	STRAIN-R-47;		
RC	MEDLINE=99241045; PubMed=10222200;		
RA	Kamitori S., Kondo S., Okuyama K., Yokota T., Shimura Y., Tonozuka T.;		
RA	Sakano Y.;		
RT	"Crystal structure of Thermoactinomyces vulgaris R-47 alpha-amylase I		
RT	(TYAIL) hydrolyzing cyclodextrins and pullulan at 2.6-A resolution.";		
RL	J. Mol. Biol. 287:907-921(1999).		
CC	-1- FUNCTION: HYDROLYSES PULLULAN EFFICIENTLY BUT ONLY A SMALL AMOUNT		
CC	OF STARCH. ENDOPHYLOLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN		
CC	PULLULAN TO FORM PANOSE. CLEAVES ALSO (1-6)-ALPHA-GLUCOSIDIC		
CC	LINKAGES TO FORM MALTOTRIOSE.		

[illegible]

Db 538 KQTVLQVPSGKTM 553

Search completed: July 31, 2002, 12:15:20  
Job time: 249 sec

---





Db 68 AERFALGACRVLAGMPAPLGCATALRGVNFVAVYSSGASASLCLFAPGDLKADRYTEE 127  
QY 105 VPLDPLMNTGNVWHVFEIG-ELHNMLYGYRDCGTFAPRCHGYLDVSNVYDPAKAVIS 163  
Db 128 VPLDPLMNTGNVWHVFEIGDQLHGMILGYRDSVAPRGGYDVSNNVYDPAKAVIS 187  
QY 164 RGEYGPARGNCCWPMQAMGIMPLPYSTFDMEGDLPURYQOKDLVITYEMHLRGFTKDDSN 223  
Db 188 RGEYGPARGNCCWPMQAMGIMPLPYSTFDMEGDLPURYQOKDLVITYEMHLRGFTKDDSN 247  
QY 224 VEHGPTGAVSKLDYKELGVNCTELMPCHEFNELEYSTSSKMFVGYSTINFEPSPT 283  
Db 248 TKHGTGYGAVSKLDYKELGVNCTELMPCHEFNELEYSTSSKMFVGYSTINFEPSPT 307  
QY 284 RYTSGLINCCRDALNEKTYREAHKRGIEYILDVFNHTAEKNENGILSPKGYDNT 343  
Db 308 RYSSSGIRBSCGALNEKAFYREAHKRGIEYIMDVFNHTAEKNENGILSPKGYDNT 367  
QY 344 YMLAPKGEFNYSGCGNTFNCNHPVROFIVDCLRYWYTEHNVHVGFRDLASIMTRGSS 403  
Db 368 YMLAPKGEFNYSGCGNTFNCNHPVROFIVDCLRYWYTEHNVHVGFRDLASIMTRGSS 427  
QY 404 LMDPVNVGAPLEGDMITTTGTPVLPPLDMISNDPILGCVKLIAMADAGLYOYGOP 463  
Db 428 LMDPVNVGAPLEGDMITTTGTPVLPPLDMISNDPILGCVKLIAMADAGLYOYGOP 487  
QY 464 HNNVSENGKTYRDIVROFIKGTDFAGFACELCGSPHLYOAGGRKPMHSINFCVACHG 523  
Db 488 HNNVSENGKTYRDIVROFIKGTDFAGFACELCGSPHLYOAGGRKPMHSINFCVACHG 547  
QY 524 FTLLADLVYNNKYNLPNGENNRDGENHNLNMGCEGEGFARLSVRLRKROMNFEVCLM 583  
Db 548 FTLLADLVYNNKYNLPNGENNRDGENHNLNMGCEGEGFARLSVRLRKROMNFEVCLM 607  
QY 584 VSOGVPMFYMGDEYHGTGKGNNTYCHDSYVNYFRMDK-EOYSELHRCCLMTKFRKCC 642  
Db 608 VSOGVPMFYMGDEYHGTGKGNNTYCHDSYVNYFRMDK-EOYSELHRCCLMTKFRKCC 667  
QY 643 EGLGLEDFPTAKRLQWGHQPKPDMSENSRFAVAFMKDEGGEIYVANTSHLPVVG 702  
Db 668 EGLGLEDFPTAKRLQWGHQPKPDMSENSRFAVAFMKDEGGEIYVANTSHLPVVG 727  
QY 703 PERAGRRPEPVVDGKPAFYDFLTDDLPDRALTIHOFSHFLYSLNLYPMLSYSSVILVLR 762  
Db 728 PERAGRRPEPVVDGKPAFYDFLTDDLPDRALTIHOFSHFLYSLNLYPMLSYSSVILVLR 787  
QY 763 DV 764  
Db 788 DV 789

RESULT 2  
Q41742 PRELIMINARY: PRT: 818 AA.  
AC Q41742: 01-NOV-1996 (Tremblrel. 01. Created)  
DT 01-NOV-1996 (Tremblrel. 01. Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19. Last annotation update)  
DE SUIP (FRAGMENT).  
GN SUGARY1.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Paulicoidae; Andropogonaceae; Zea.  
OX NCBI\_taxid=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-ENDOSPERM.  
RX MEDLINE=95290995; PubMed=7773016;  
RA James M.G., Robertson D.S., Myers A.M.;  
RT "Characterization of the maize gene sugary1, a determinant of starch  
composition in kernels.";  
RL Plant Cell 7:417-429(1995).

DR EMBL: U18908; AAA91298.1; -  
DR HSSP: P10342; 1BR2.  
DR InterPro: IPR000461; AlphaAmylase.  
DR InterPro: IPR004193; Isoamylase\_N.  
DR Pfam: PF00128; alpha-amyase; 1.  
DR Pfam: PF02922; Isoamylase\_N; 1.  
FT NON\_TER  
SQ SEQUENCE 818 AA; 91535 MW; 4779CA8508B1AFED CRC64;  
  
Query Match 83.7%; Score 3527.5; DB 10; Length 818;  
Best Local Similarity 83.0%; Pred. No. 2e-273;  
Matches 649; Conservative 44; Mismatches 70; Indels 19; Gaps 6;  
  
QY 1 SGPAPLR-----RW-----RNATAGKGVGY-----CAAVEATVVEDEGEDEY 44  
Db 38 SSRLPLAVPAGRWAGVGRVGRN-VAGLGRGRLSLHAAARPAVEVQAEEDDDDEY 96  
QY 45 AEDRYALGACRVLAGMPAPLGCATALRGVNFVAVYSSGATAAALCFEPEDLKADRYTEE 104  
Db 97 AERFALGACRVLAGMPAPLGCATALRGVNFVAVYSSGASASLCLFAPGDLKADRYTEE 156  
QY 105 VPLDPLMNTGNVWHVFEIG-ELHNMLYGYRDCGTFAPRCHGYLDVSNVYDPAKAVIS 163  
Db 157 VPLDPLMNTGNVWHVFEIGDQLHGMILGYRDSVAPRGGYDVSNNVYDPAKAVIS 216  
QY 164 RGEYGPARGNCCWPMQAMGIMPLPYSTFDMEGDLPURYQOKDLVITYEMHLRGFTKDDSN 223  
Db 217 RGEYGPARGNCCWPMQAMGIMPLPYSTFDMEGDLPURYQOKDLVITYEMHLRGFTKDDSN 276  
QY 224 VEHGPTGAVSKLDYKELGVNCTELMPCHEFNELEYSTSSKMFVGYSTINFEPSPT 283  
Db 277 TKHGTGYGAVSKLDYKELGVNCTELMPCHEFNELEYSTSSKMFVGYSTINFEPSPT 336  
QY 284 RYTSGLINCCRDALNEKTYREAHKRGIEYILDVFNHTAEKNENGILSPKGYDNT 343  
Db 337 RYSSSGIRBSCGALNEKAFYREAHKRGIEYIMDVFNHTAEKNENGILSPKGYDNT 396  
QY 344 YMLAPKGEFNYSGCGNTFNCNHPVROFIVDCLRYWYTEHNVHVGFRDLASIMTRGSS 403  
Db 397 YMLAPKGEFNYSGCGNTFNCNHPVROFIVDCLRYWYTEHNVHVGFRDLASIMTRGSS 456  
QY 404 LMDPVNVGAPLEGDMITTTGTPVLPPLDMISNDPILGCVKLIAMADAGLYOYGOP 463  
Db 457 LMDPVNVGAPLEGDMITTTGTPVLPPLDMISNDPILGCVKLIAMADAGLYOYGOP 516  
QY 464 HNNVSENGKTYRDIVROFIKGTDFAGFACELCGSPHLYOAGGRKPMHSINFCVACHG 523  
Db 517 HNNVSENGKTYRDIVROFIKGTDFAGFACELCGSPHLYOAGGRKPMHSINFCVACHG 576  
QY 524 FTLLADLVYNNKYNLPNGENNRDGENHNLNMGCEGEGFARLSVRLRKROMNFEVCLM 583  
Db 577 FTLLADLVYNNKYNLPNGENNRDGENHNLNMGCEGEGFARLSVRLRKROMNFEVCLM 636  
QY 584 VSOGVPMFYMGDEYHGTGKGNNTYCHDSYVNYFRMDK-EOYSELHRCCLMTKFRKCC 642  
Db 637 VSOGVPMFYMGDEYHGTGKGNNTYCHDSYVNYFRMDK-EOYSELHRCCLMTKFRKCC 696  
QY 643 EGLGLEDFPTAKRLQWGHQPKPDMSENSRFAVAFMKDEGGEIYVANTSHLPVVG 702  
Db 697 EGLGLEDFPTAKRLQWGHQPKPDMSENSRFAVAFMKDEGGEIYVANTSHLPVVG 756  
QY 703 PERAGRRPEPVVDGKPAFYDFLTDDLPDRALTIHOFSHFLYSLNLYPMLSYSSVILVLR 762  
Db 757 PERAGRRPEPVVDGKPAFYDFLTDDLPDRALTIHOFSHFLYSLNLYPMLSYSSVILVLR 816  
QY 763 DV 764  
Db 817 DV 818

RESULT 3  
080403

DT	080403	PRELIMINARY;	PRT;	733	AA.
AC	080403;				
DT	01-NOV-1998	(TRMBUREL. 08, Created)			
DT	01-NOV-1998	(TRMBUREL. 08, last sequence update)			
DT	01-DEC-2001	(TEMBUREL. 19, last annotation update)			
DE	ISOMYLASE (FRAGMENT).				
OS	Oryza sativa (Rice).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Eihartoideae; Oryzaceae; Oryza.				
OX	NCBI_TaxID=4530;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. FUJIIKIRI; TISSUE=ENDOSPERM;				
RX	MEDLINE=99266133; PubMed=10333501;				
RA	Fujita N., Kudo A., Francisco P.B., Nakakita M., Harada K., Minaka N.,				
RA	Nakamura Y.;				
RT	"Purification, characterization, and cDNA structure of isomylase from				
RT	developing endosperm of rice."				
RL	Planta 208:283-293(1999).				
DR	EMBL; AB015615; BAA29041.1; -.				
DR	HSSP; P10342; 1BF2.				
DR	InterPro: IPR000461; Alpha_amyase.				
DR	InterPro: IPR004193; isomylase_N.				
DR	Pfam: PF00128; alpha-amyase.1.				
DR	Pfam: PF02922; isomylase_N.1.				
FT	NON_TER				
FT	1				
FT	1				
SO	SEQUENCE	733	AA;	82119	MM; FC864DDDF0E3D035 CRC64;

[illegible][illegible]

SEQUENCE	569 AA	64020 MW	5E5B8B12427D435A	CRC64
Q9XFG6	PRELIMINARY	PRT	569	AA
ID	Q9XFG6			
AC	Q9XFG6			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	ISOMYLASE 1 (FRAGMENT).			
OS	Hordeum vulgare (Barley).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.			
OX	NCBI_TaxID=4513;			
NP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. BOMI;			
RX	MEDLINE=99364538; PubMed=104337827;			
RA	Sun C., Sathish P., Ahlendsberg S., Jansson C.;			
RT	"Analyses of isomylase gene activity in wild-type barley indicate its involvement in starch synthesis."			
RL	Plant Mol. Biol. 40:431-443(1999).			
DR	EMBL: AF142589; AAD33889.1; .			
DR	HSSP: P10342; 1BF2.			
DR	InterPro: IPR00461; Alpha_amylose.			
DR	InterPro: IPR004193; isomylase_N.			
DR	Pfam: PF00128; alpha-amylose; 1.			
DR	Pfam: PF02922; isomylase_N; 1.			
FT	NON_TER			
FT	NON_TER			

Query Match	71.2%	Score 3001	DB 10	Length 569
Best Local Similarity	94.7%	Pred. No. 1.5e-231		
Matches 541	Conservative 12	Mismatches 16	Indels 2	Gaps 2
OY	1 GCGACRVLAGMAPAPGATPLAGGVNFAYVSGGATAAALCTFPEDLKAAKRVTEEYVLDPL	110		
DB	1 GCGACRVLAGMRA - AGATPLAGGVNFAYIS - AEPALALCTFPEDLKAAKRVSEYVLDPL	58		
OY	111 NMRGNVWHVFTEGELHNKLGYGDFDGTFFAPHCGHYLDVSNVVDPYAKAVISRGEGVP	170		
DB	59 NMRGTDVWHVFTEGELHGMLYGDFDGTFFAPHCGHYFDVSNVVDPYAKAVISREEGVP	118		
OY	121 ARGNNCQPMQAMNITPLSTPDMEDDPLRYQOKOLYTEMHLRGTKKDDSNVHEPGTF	230		
DB	119 ARGNNCQPMQAMNITPLSTPDMEDDPLRYQOKOLYTEMHLRGTKKDDSNVHEPGTF	178		
OY	231 IGAVSKLDYLKELGVNCTELMPCHGFENLEYSTSSKKNFMFGYSTINFPSPMTRYSGGI	290		
DB	179 IGAVSKLDYLKELGVNCTELMPCHGFENLEYSTSSKKNFMFGYSTINFPSPMTRYSGGI	238		
OY	291 KNCGDALINEKTFPVRBAHKRGIEVILDVFNHTAEGNENGPILSPKGVNDTTYIWLAKR	350		
DB	239 KNCGRDGINEKTFPVRBAHKRGIEVILDVFNHTAEGNENGPILSPKGVNDTTYIWLAKR	298		
OY	351 GEFYVYSGGNFENCNHPPVROFIYDCLRYWTEEMHNVGSEFRPDLASIMTRGSSLMDPPVY	410		
DB	299 GEFYVYSGGNFENCNHPPVROFIYDCLRYWTEEMHNVGSEFRPDLASIMTRGSSLMDPPVY	358		
OY	411 YGAPLEGDMITTGRLPYTPPLDLMSINPDLIGVXKLAEAMAGGLXYVGQGPFRHNVWSE	470		
DB	359 YGAPLEGDMITTGRLPYTPPLDLMSINPDLIGVXKLAEAMAGGLXYVGQGPFRHNVWSE	418		

Oy	471	MNGKRDIVROFIKGTDEPAGCFACCLGSPHLYOAGCKPMHSINFCADHGFTLADLV	530
Dd	419	MNGKRDIVROFIKGTDEPAGCFACCLGSPHLYOAGCKPMHSINFCADHGFTLADLV	478
Oy	531	TYNKKYNNLPNGENNNDGEGENHNLSNMGCEGEFAPRLSVKLRRQRNRNFVCLMVSQGYPM	590
Dd	479	TYHNKYNNLPNGEDNNDGEGEHNHNSNMKGEGEFARBSVLRRLRQRNRNPFVCLMVSQGYPI	538
Oy	591	FYMGGDEYGHTKGNNNTYCSDSYVNYFRWDK	621
Dd	539	FYMGGDEYGHTKGNNNTYCSDSYVNYFRWEK	569
RESULT	5		
ID	004196	PRELIMINARY:	PRT: 783 AA.
AC	004196:		
Dt	01-JUL-1997	(TREMBLrel. 04, Created)	
Dt	01-JUL-1997	(TREMBLrel. 04, Last sequence update)	
Dt	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
De	POTATIVE ISOAMYLASE.		
Gn	AT2G39930.		
Oc	Arabidopsis thaliana (Mouse-ear cress).		
Oc	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
Oc	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
Ox	eurosidis II; Brassicales; Brassicaceae; Arabidopsis.		
Ncbi	_taxid=3702;		
Rn	[1]		
Rp	SEQUENCE FROM N.A.		
Rc	STRAIN=CV. COLUMBIA;		
Rx	MEDLINE=20083487; PubMed=10617197;		
Ra	Ltn X., Keul S., Rounsley S.D., Shea T.P., Bentlo M.-T., Town C.D.,		
Ra	Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,		
Ra	Buell C.R., Ketchum K.A., Lee J.J., Romling C.M., Koo H., Motil K.S.,		
Ra	Cronin L.A., Sherran L.M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,		
Ra	Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,		
Ra	Copenhaver G.P., Preuss D., Niernm J.C., White O., Eisen J.A.,		
Ra	Salzberg S.L., Fraser C.M., Venter J.C.:		
Rt	"Sequence and analysis of chromosome 2 of the plant Arabidopsis		
Rt	thaliana."		
Rl	Nature 402:761-768(1999).		
Rn	[2]		
Rp	SEQUENCE FROM N.A.		
Rc	STRAIN=CV. COLUMBIA;		
Ra	Ltn X.:		
Rl	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
Dr	EMBL: AF002109; AAB95278.1; .		
Dr	HSSB: P10342; IBF2.		
Dr	InterPro: IPRO00461; Alpha_amyase.		
Dr	InterPro: IPRO00493; IsoAmyLase_N.		
Dr	pfam: PF02922; alpha-amyase; 1.		
Dr	pfam: PF02922; isoAmyLase_N; 1.		
Dq	SEQUENCE 783 AA: 89480 MW; 638A6F9088CDB896 CRC64;		
Query Match	68.7%: Score 2895.5; DB 10; Length 783;		
Best Local Similarity	69.3%: Pred. No. 7e-223;		
Matches 516: Conservative	84; Mismatches 128; Indels 17; Gaps		
Oy	28	VEAAARKVEDEGE---DPPVAEDRYALGACGRVLGMPAPLGPATAGLVNFAVYSGA	83
Dd	42	ISAKORSRSEANINAVEERPLKSDFRID-----GLPSFGPGYVRDDGVNFSYSTS	95
Oy	84	TAAALCLFPEDLKADRYTEEVFLDPLMRKTGNVHVLEEGELHNLXYTRDDGTFAPPC	143
Dd	96	VSAITICLISLDLRONKRYTEEIQLDPESNRKTHVMVFIRLGDFKMLYGVRDGKRSPBE	155
Oy	144	GHYLDVSNVVVPYKAVYSRGEGYPARGNNCMQPOMAMIPLPYSTFMEGDLPLRYQ	203
Dd	156	GHYDSSNLLDLPYKAIIISRDEFYGLGPDNDNCWPMACMVPRREEFFMEGDMLHLKPQ	215
Oy	204	KDLVIYEMLRGFETKHDSNVHEPTGIFAVSKLDYKLDELGVNCTELMDCHENELEYST	263

Dd		216	KDLVYEMHVGFTPTTHESKKTIEPCTYGVAKENCDLHKELGNCLEMLRPHCEHNLELEY	275
Qy	264	SSS----	KMFNGYSTINFSFPMTRYTSGGIKCGRDAINEFKTFVBRAHRGIEVILD	318
Dd	276	YNTILGDHRVNFWGSTIGCFSPMIRYASASNNPAGRAINDEFKLTYEAKRGLEVIMD		335
Qy	319	VVENITAGGENEGPLSTKCYQNDNTYYLMARKEFYNTSGCGNFPCNCHPPVROGIYUCL		378
Dd	336	VVLNHTJAGGENEGPIEFSEFRGYDNYSYYMLAPKGEPYNTSGCGNFPCNCHPPVROPIIDCL		395
Qy	379	RWVWEHVHDGRRPFLASTMTFGSSLTAMPVVNYCYGARPEGDMITGTPTLYTPPLDIMISND		438
Dd	386	RWTWTEMHVDGRRPFLGSIMSRSLLMPAANYGADVBGDDLTITGTPTLSCPVYIDMISND		455
Qy	439	PILGGVKLIAEAMDAGGLYOYGQPEFHMMVMSEMNKKYDIYROIQTIGTDFAGFAECIC		498
Dd	456	PILRGVKLIAEAMDAGGLYOGMEPFHWIWSMNCKFPDYVRQPIFKGIDGFSGAFAELC		515
Qy	499	GSPHLTYOAGRRKPWHISINFCVCAHDEFTLADLVTKKYKNLPNGENNRDGENHNLSWNGE		558
Dd	516	GSPNLTYO--GGRKPMHSINFICAHDEFTLADLVTKYNKNNLNLAEGENNENHNYSMNCGE		574
Qy	559	BEEFARLSYKRRLRKQOMNEFFYCLAWISOGVPWFYMGDEYGHTKGKNNTMYCHDSVNVNR		618
Dd	575	EEDFASISYKRLRKQRQMREFYSLVNSOGVPMIYMGDEYGHTRKGGNNNTYCHDNTMYNR		634
Qy	619	WDKKSO-YSELRRPCCMLTKPRKECEGLEGDFPTAKRLQMHSHQPGKPDMSSENSRFVAE		677
Dd	635	WDKKEAHSDFFRCFRILIKFRDECESLGANDFPRAKRLQHGLAPELPNMSETSRFAAF		694
Qy	678	SMKDERQGEIYVAFMNTSHLPAVELPERAGRREBVDVTGRKAPAYDTLDLPBALRTIH		737
Dd	695	SILVDYSKKEIYVAFMNTSHLATLVSLPNBPGRYRMFBFDVTSKPSPYDCITPDLPERETAMK		754
Qy	738	QSFHLSNLTYPMLSYSVIYLRP	762	
Dd	755	QYRHFLDANVYPMLSYSTIILLSP	779	
 RESULT 6 Q9XFG7 PRELIMINARY: PRT: 327 AA.				
ID	Q9XFG7	AC	Q9XFG7	
Dt	01-NOV-1999	(TREMBLrel. 12, Created)		
Dt	01-NOV-1999	(TREMBLrel. 12, Last sequence update)		
Dt	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	ISOAMYLAZE 1 (FRAGMENT).			
OS	Trilicium aestivum (wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spemmatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;			
OC	Triliceae; Trilicum.			
Ox	NCBI_TaxID=4565;			
Rn	[1]			
Rp	SEQUENCE FROM N.A.			
Rx	MEDLINE=99364538;	PubMed=10437827;		
Rx	Sun C., Sathish P., Ahlandsberg S., Jansson C.;			
RA	'Analyses of isoamylase gene activity in wild-type barley indicate its			
RT	involvement in starch synthesis";			
RL	Plant Mol. Biol. 40:431-443(1999).			
DR	EMBL; AF142590; AAD33890.1; -.			
DR	HSSP; P10342; IBF2.			
DR	InterPro: IPR000461; Alpha-amylase.			
DR	pfam: PF00128; alpha-amylase; 1.			
DR	PRINTS: PR00110; ALPHAMYLASE.			
FT	NON_TER	1		
FT	NON_TER	327		
SEQ	SEQUENCE	327 AA; 36876 MW; 9474EB0BEF4F07B9 CRC64;		

Qy	196	DLLRYPKKDLVLEYEMHKGRTKDDSSNVEHPGFFICAVSKLDYIAKELGVNCEIEMPCH	255
Db	1	DLLRYPKKDLVLEYEMHKGRTKDDSSNVEHPGFFICAVSKLDYIAKELGVNCEIEMPCH	60
Qy	256	FNELEYSRSSSKMNMWGSTINFEPSPMTRYSGGIRKNGCDAINEFPTFYREAHKRGIEV	315
Db	61	FNELEYSRSSSKMNMWGSTINFEPSPMTRYSGGIRKNGCDAINEFPTFYREAHKRGIEV	120
Qy	316	ILDVVFENHTAAGNENGPIPLSEKGVNDNTYYMLAPKGEFYNVSGCGNFNCNHHVPRQFIY	375
Db	121	ILDVVFENHTAAGNENGPIPLSRGVNDNTYYMLAPKGEFYNVSGCGNFNCNHHVPRQFIY	180
Qy	376	DCLRWYTEMHVDGFRDLASIMTRGSSLMDPVNVYCAPIEGDMITTGPLVTPPLIDMI	435
Db	181	DCLRWYTEMHVDGFRDLASIMTRGSSLMDPVNVYCAPIEGDMITTGPLVTPPLIDMI	240
Qy	436	SNDPLIGGVKILAEKMDAGCIYQVGQRPNNVNMSENNKTRIDYIROPFIKGTGDFAGGFAE	495
Db	241	SNDPLIGGVKILAEKMDAGCIYQVGQRPNNVNMSENNKTRIDYIROPFIKGTGDFAGGFAE	300
Qy	496	CLCGSPHLIYQAGGRKPWHSINFEVCAHD	522
Db	301	CLCGSPHLIYQAGGRKPWHSINFEVCAHD	327

RESULT	7
P73608	
ID	P73608
AC	P73608
DT	01-FEB-1997 (TrEMBLrel. 02, Created)
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	GLYCOSYL OPERON PROTEIN GLGX.
GN	GLGX OR SLR1857.
OS	Synechocystis sp. (strain PCC 6803).
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX	NCBI_TaxId=1148;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97061201; PubMed=8905231;
RA	Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA	Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,
RA	Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA	Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA	Tabata S.;
RT	"Sequence analysis of the gene of the unicellular cyanobacterium
RT	Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT	entire genome and assignment of potential protein-coding regions."
RL	DNA Res. 3:109-136(1996).
DR	EMBL; D90908; BAA17652.1; ..
DR	HSSP; P10342; 1BF2.
DR	InterPro; IPR000461; Alpha_amyase.
DR	InterPro; IPR004193; Isoamylase_N.
DR	Pfam; PF00128; alpha-amyase_1.
DR	Pfam; PF02922; Isoamylase_N; 1.
DR	Complete proteome.
QO	SEQUENCE 707 AA; 79895 MW; CFA27B3C86318DB1 CRC64;

[illegible]

```

RESULT      8
ID           09M055      PRELIMINARY;      PRT;      702 AA.
AC           09M055;
DT           01-OCT-2000 (TrEMBLrel. 15, Created)
DT           01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT           01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE           ISOAMYLASE-LIKE PROTEIN.
OS           AtG090920.
OS           Arabidopsis thaliana (Mouse-ear cress).
OC           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC           Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX           NCBI_TaxID=3702;
RN           [1]
RP           SEQUENCE FROM N.A.
RA           Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA           Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA           Shekhar M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA           Mewes H.W., Lemcke K., Meyer K.F.X.;
RL           Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN           [2]
RP           SEQUENCE FROM N.A.
RA           EU Arabidopsis sequencing project;
RL           Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR           EMBL: AL161513; CAB78026.1; -.
DR           HSSP: P10342; 1BP2.
DR           InterPro: IPR000461; Alpha-amylase.
DR           Pfam: PF00128; Alpha-amylase; 1.
DR           SEQUENCE 702 AA: 79558 MW: 264f66c3c30cf942 CRC64:

```





Db	388	EMNGPYRDNVAFKLGDNQLISTFASRISGSODITYPHG--SPTNSINVSCHDGFLLCOT	445
Oy	530	VTYKKYVLPNGENNROGENNHNLSNCGEGEFARLSYKRLKRRQMRNFVCLAMSQV	589
Db	446	VTYNKHNKHEANGEDNRDGDANYSYNFGTEGKTEDPGILEVERERQIRNFELTLMSQGI	505
Oy	590	MEYMGDEGHTKRGNNNTYCHDSYVNYFPMWKKEDQYSELHRCCLMTKRKECEGGLD	649
Db	506	MIQSDDEVAHTAEAGNNNRALDLSNANYPELMDLTAKPTLMHFLCDLAIARKKYKLTENRG	565
Oy	650	FPTAKRLQMHGHPCKPDMSENSERFVAFSMKDEROGEIYVAFNTSHLPVLELPE	704
Db	566	FLSKKEISWVDAMGMPMTWRPBN-FLAKRIKSPK-AHYVAFHVGADQDLATLPK	618
RESULT 12			
ID	084046	PRELIMINARY: PRT: 666 AA.	
AC	084046		
DT	01-NOV-1998	(TREMBLrel. 08, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	GLYCOSYL HYDROLASE (DEBRANCHING).		
GN	GGX OR CT042.		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=813;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-D/UM-3/CX.		
RX	MEDLINE=99000809; PubMed=9784136;		
RA	Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,		
RA	Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koehn E.V.,		
RA	Davis R.W.;		
RT	"Genome sequence of an obligate intracellular pathogen of humans:		
RT	Chlamydia trachomatis."		
RL	Science 282:754-759(1998).		
DR	EMBL: AE001278; AAC67632.1; .		
DR	HSSP; P10342; 1BF2.		
DR	InterPro: IPR004193; Isoamylase_N.		
DR	InterPro: IPR004193; Isoamylase_N.		
DR	Pfam: PF00128; alpha-amylose; 1.		
DR	Pfam: PF02922; isoamylase_N; 1.		
KM	HydroLase; Complete proteome.		
SO	SEQUENCE 666 AA; 76076 MW; ES8F03AB064FECE7F CRC64;		
Query Match 30.9%; Score 1300; DB 16; Length 666;			
Best Local Similarity 39.8%; Pred. No. 2.6e-95;			
Matches 289; Conservative 102; Mismatches 241; Indels 94; Gaps 20;			
Oy	57	VLGAPAPLIGATL-AGGVNFAVSGAGTAALCLFETPEDLKAADRYTE--EYPLDPLMNR	113
Db	6	VRSTIPLPLGAKKLSADRRFRSLFSSQAQVTLVLDL-----LSETHETPLSTDR	58
Oy	114	TGNVNHVTEGELHNMLTYGRFDGT-----FAPHCGHYLDVSNVVDYAKAVISRG	167
Db	59	TGAIWHIETIAGISSMSYAYKLRGDLSSQKRA-----TDSYIADPYSKNITSOLF	110
Oy	168	GVPARGNCCWPMAGMIDPLPSTEDWEGDLPLRYOKDLVIEMLRGFTKIDSSVNEHP	227
Db	111	GSPKOEK-----VAFSYLKHEDFPEMGCTPLHLRKENFIEMHVRSPTRPDSSQVSHR	165
Oy	228	GTFPGAVSKLDYLRKLGNGVCIETELMCPHENELYSTSSKM-----NEMXSTINFPSPMT	283
Db	166	GTFPGIETIKIDHLKQGVNAVELLTFEEDETVHPKNDPRLHLCNLYWSSVNFECPSR	225
Oy	284	RYTSGGINCGRDAINERKTEVREARHKKGIEYILDVFNHNTAEGNENGPILSFKGVDNPT	343
Db	226	RYTGA-DECA--PARREKTLVKAHARGIEVYILDVFNHNTGFEIGSCL-----PMDLES	279
Oy	344	YTM LAPKEGYNYSGCGNTFCNBNPVRQFIVDCLRVWYTEMHNVDGFRDLASIMTRGSS	403

```

Db      280  YYMVDHGDLDLNFSCSGMTVTNTPPTLLKMLDALRYWQEKHHVDFGRDLASVFSR--- 336
QY      404  LMDPVNVTGAPTEGDMITTPGTPVLPTRPLDIMSNDPLIGCVLLIEMAPAGGLYQVGF 463
Db      337  -Dp-----OGVPLPTRLIQAISDSILSEFTKLIAEPWDAGGLYQLGHP 380
QY      464  HMNV-WSEMNKGYRDLVYQFIKGTGFGAGCFACDLCGSHLVQAGRKRWHSINFCAD 522
Db      381  SISTWSEMNKGYRDLVYQFVLAFLNGDAHQVSSFSRISGSHDIPNG--KPTNSINYSCHD 438
QY      523  GFTLADLVTYNNKYRLPNGENNRDGENNHLNNNGCEGEPFARLSYRLKRPQMRNFVCL 582
Db      439  GFTLIDVYAIYNDKHNDEKNEYNRDGTSAHYSTNFCGEGTTPPTICALERQKNEFLAL 498
QY      583  MVSQGVPMFYMGDEYGHTRKGNMNTYCHDSYVYFRMDKKEQYSELHRECCMLTKRKEC 642
Db      499  FLSQGIPIAQSDGYHTAYGNMNMCLDPTKINFLMDRLARKELEFSFLQVIALRKAY 558
QY      643  ECLGEDEFPYAKRLQWGHQPKPWSSEKRVAAFSMKDERGEIYAANTSHLPAYVEL 702
Db      559  TELFTMTSLSEDTITMLWTKGSPREMGAD-HYLAELK-HLNYSLPFAVYSGHERIEISL 616
QY      703  PE--RAGRMRPEVVD--NGKPAFYFLNDDLPBALTIHQFSHFLXSNLYPMNS---YSS 755
Db      617  PKPKHEHLATEKIVDSTTG-----FSSQLISPKLSLEPYSS 652
QY      756  VILVLR 761
Db      653  LVAIAR 658

RESULT  13
Q9RXP5  Q9RXP5  PRELIMINARY;  PRT;  720 AA.
AC  Q9RXP5.
DT  01-MAY-2000 (TREMBLrel. 13, Created)
DT  01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  GLYCOGEN OPERON PROTEIN GLGX.
GN  DR0264.
OS  Deinococcus radiodurans.
OC  Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX  NCBI_TaxID=1299;
   [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=RI.
RX  MEDLINE=20036896; PubMed=10567266;
RA  White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA  Dossan R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA  Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA  Matkarova K.S., Aravind L., McDonald L., Uutterback T., Zalewski C.,
RA  Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA  Fraser C.M.;
RT  "Genome sequence of the radioresistant bacterium Deinococcus
RT  radiodurans R1."
RL  Science 286:1571-1577(1999).
DR  EMBL; AE001888; AAF09848.1; .
DR  HSSP; P10342; 1BF2.
DR  TIGR; DR0264; .
DR  InterPro; IPR000461; Alpha_amlase.
DR  InterPro; IPR004193; Isoamylase_N.
DR  Pfam; PF00128; alpha-amylose; 1.
DR  Pfam; PF02922; Isoamylase_N; 1.
KW  Complete proteome.
SQ  SEQUENCE 720 AA; 80613 MM; DE23118CBBE3E7157 CRC64;

```



```

QY 60 GMRPRLGATAGVNFVAVISGATAAALCLFTEPBDLKADRVTEEVPDLPLMNRGTQVMH 119
DB 11 GSEPLGATATDCKGTNFALYSENATGYELCLDAEGHET-----REFPLEQAFVWH 62
QY 120 VFELEGELHNLGYRFRFGTAPRHGHLDVSNVYVDPYAKAV-----ISREKYVPARGN 174
DB 63 GYLPGLGPGRGYRGVHGEVAPENKGLRFPNPNVLLDPYAKALDGTEDGRGVGVAGE 122
QY 175 NCMPPOM-----AGMPLPYSTFDMEDGLPLRYPOKDLVYEMHLRGFT-KHDSNV 224
DB 123 DD-SQMGEEGRGAPGLGVDP--MFWVGDQKPGIFPHOSVITEAHVKGLTMTHPDVE 179
QY 225 EHPGTIGAVSK--LDYLKELGVNLCIEMPCHEFNELEYSTSSKMFQWSTINFSPM 282
DB 180 ELRGTYAGVATPAIDLTLDLGTALIEFLVHQHVDDPFLDKLTVMGWTINFPARD 239
QY 283 TRTSGGKRCGRDAINEFTFYREAHKRGLIEVLDVFNHTAGENGPILSKGVNT 342
DB 240 VRTSARRGKPNPSGAVPEFNMVRLHDAGIEVLDVYVHTAGNMGPTMSFKIDNP 299
QY 343 TYVMLAPKGE--FYNSGCGNTFNCNHPVYRQFIVDCLRYVWEMHVDGPRDLASTMR 400
DB 300 TYRLVADDOREFYFDYTGTSNLSNVRHPOTLQIMSLRYWTEMHVDGFRFDLASTLAR 359
QY 401 GSSLMDPVNVYGAPIEGDMITTTGTPPLPLIDMISNDPILGKYLAEAMADG-GLYOV 459
DB 360 GLHEVYDQLSGF-----FTIIHQDPLISQVKLAEFMDVDEGGSTOV 399
QY 460 GCFPHNHNWSEWNGKYRDIYKROFIKTDGAGFAECLGSPHLVYQAGRRPMHSINFC 519
DB 400 GNEP--VNNAEVNGIYRDMRSFWMKGGGLASEIGYRITSSDLYEPNGKPYASINFT 457
QY 520 AHDGFTLADLYTNKKYKLPNGENNRDGENHNLWNGEGEFARLSVKLRKQMANFF 579
DB 458 AHDGFTLRDSVYTEQHNNEANGNDGNHNTIMNCGVGGPTDDEINRLRQOQMNFL 517
QY 580 VCLMVSQGVPMFYMGDEYGHHTKGNNTYCHDSYVNFYRMDK-----KEQ 624
DB 518 ATLLGQGTFRMLGDEFEPTGCGNNNAQCQDDISMYDEKXDELLATRLILARKA 577
QY 625 YSELHFRCCMLKFRKCEGGLG--EDFPYAKRLQNHGHPGRPDSENSRFAVFSM-- 679
DB 578 HPSLHR-----RKFFAGRNIRGEDVRIYVLRFDGAEEMDEDMN-INQTSLSGFLA 628
QY 680 -----KDERGCELYVAFNTHSLRAYVELPERAG--RRPEPVYDT 716
DB 629 GDGLADVAEGKPLTDDHLLLSSTSYVDLPKMPDLGGCGEWDLLDT 677

```

RESULT 14

Q9PK26 PRELIMINARY: PRT: 666 AA.

AC Q9PK26: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

GN GLYCOSYL HYDROLASE FAMILY PROTEIN.

OS Chlamydia muridarum

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

ON NCBI\_TaxID=83560;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MOPN / NIGG

RX MEDLINE=20150255; PubMed=10684935.

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Glinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,

RA Eisen J., Fraser C.M.

RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia

RT pneumoniae A39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

```

DR EMBL: AE002298; AAF39177.1; -
DR HSSP: P10342; 18F2.
DR TIGR: TC0312; -.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
DR Complete proteome.
SQ SEQUENCE 666 AA; 76232 MW; B7D9C61F0ED1B684 CRC64;

Query Match      30.0%; Score 1264; DB 16; Length 666;
Best Local Similarity 38.4%; Pred. No. 2e-92;
Matches 282; Conservative 105; Mismatches 245; Indels 102; Gaps 20;

QY 53 GACVYLAGMRPRLGATAGG-VNFVYSGGATAAALCLFTEPBDLKADRVTEEVPDLPLM 111
DB 2 GSLSHLTLPLPLGAOKLSDRYRFSLVAPRAQOYILVLDPFSS-----EIHETPLSSTD 56
QY 112 NRTGNVHVFTEGELHNLGYR--DGTAPRHGHLDVSNVYVDPYAKAVISRGEYG 168
DB 57 HRTCAIMHIEISGISNENSTAYKLRQSDSAF-PN---FSTNAYIADPYSKNIFS----- 106
QY 169 VPARGNCPQMGAMTLP-----YSTFDMEDGLPLRYPOKDLVYEMHLRGFTKH 219
DB 107 -----POLFSSSKQPNDYTFSTYKQEDPDMEDGTPPLPLKENTFYTEMHRSFTOD 157
QY 220 DSSVNEHPGTIFGAVSKLDYLKELGVNLCIEMPCHEFNELEYSTSSKMF--NFWGYST 275
DB 158 PSSQVTHRGTFELGIEIKIDHLKGLVNAVVELLPLEFDEDTIHPPKKNDFPLCLNMGWISS 217
QY 276 INFSPMTRYSGGIKKCGRDAINEFTFYREAHKRGLIEVLDVFNHTAGENGPILS 335
DB 218 INFSPMTRYS--GTDCA--PAREFTLVKTLHRAIEVLDVFNHTAGENGPILS 272
QY 336 FKGYDNTTYMLAPKGEFYVNSGCGNTFNCNHPVYRQFIVDCLRYWTEMHVDGPRDLA 395
DB 273 -PWLDESITYWVNDGSLNLFSGCGNTVNTTPAIRKMLDALRYWQEMHVDGPRDLA 331
QY 396 SINTRGSSLMDPVNVYGAPIEGDMITTTGTPPLPLIDMISNDPILGKYLAEAMADG 455
DB 332 AVFSRD-----LQGVRSULPILQALISSDSILETKLIAEPMDAGG 372
QY 456 LYQVGGFPHNWN--WSEWNGKYRDIYKROFIKTDGAGFAECLGSPHLVYQAGRRPMHS 514
DB 373 LYQVGGFPHNWN--WSEWNGKYRDIYKROFIKTDGAGFAECLGSPHLVYQAGRRPMHS 514
QY 431 INFVCAHDGFTLADLYTNKKYKLPNGENNRDGENHNLWNGEGEFARLSVKLRKQ 574
DB 431 INYICSHDGTLYDSVAYNKHNEENNRDGTSAVSYNFCGEGETDPNIOQLRERQ 490
QY 575 MRNFEVCLMVSQGVPMFYMGDEYGHHTKGNNTYCHDSYVNFYRMDKEQYSELHFRCL 634
DB 491 MKNFFLALFLSQGILPMIKSDEYGHHTAYGNNNMWCDLTKLNHLMDLAEKKEFFSLCQ 550
QY 635 MTFRECEBGLGDEDPYAKRLQNHGHPGRPDSENSRFAVFSMDEROGEITYVAFNTS 694
DB 551 IITLRTTHALFNTNPLSETTITWLSOGILPRBMTDP-HYLAELKHPNY-SLFIAYSG 608
QY 695 HLPAYVELPE--RAGRMEPVVD--TGKRAPYVFLDDDLPRALTLTHQSFHFLXSNLYPM 750
DB 609 NEXIELALFKLQEHALAEKIYDSTTG-----FFSQILSPK 644
QY 751 LS---YSSVILVLR 761
DB 645 LSLPEYSSLVATSR 658

```

RESULT 15

Q9KY03 PRELIMINARY: PRT: 715 AA.

AC Q9KY03: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PUTATIVE GLYCOSYL DEBRANCHING ENZYME.  
GN GLGX.  
OS Streptomyces coelicolor.  
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;  
OC Actinomycetales: Streptomycinae: Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Seeger K.J., Harris D.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL: AL356932; CAB92884.1; -.  
DR HSSP: P10342; 1BF2.  
DR InterPro: IPR000461; Alpha\_acylase.  
DR InterPro: IPR004193; Isoamylase\_N.  
DR Pfam: PF00128; alpha-amylose; 1.  
DR Pfam: PF02922; Isoamylase\_N; 1.  
SQ SEQUENCE 715 AA; 80520 MW; 364C09DE6395D2CE CRC64;

Query Match 29.7%; Score 1251; DB 2; Length 715;

Best Local Similarity 38.08; Pred. No. 2.5e-91; Matches 290; Conservative 92; Mismatches 259; Indels 122; Gaps 20;

QY 57 VLAGMPALGATATAGAGVFAVYSGATTAALCLFTPEDLKADRYTEVPDPLMNRGN 116  
DB 13 VMSGHPPLCAAYDGGTGNFALFSEVNERVDVL-VDDGNHSTY-----PLPDVDGF 64  
QY 117 VMHVEIGELHNLGYRFDGTFAPHCGLYDVSNVVDYPAKAV-----ISRGEY 167  
DB 65 VMHCYLPVGFGQRYGYRVHGMARPAVGHRCNPAKLLDPYTRAVDGLVDNHAFLFERAR 124  
QY 168 GVPARGNCCPMQAGMPLPYSTFEDWEGDPLRYPOKDLVYEMHLRGFTK-HDSSNVEH 226  
DB 125 GKADPGDSAGHTMLGVYTDPF--FDWGDDRPFRPRYSVLEYEAHVRLGSKTHDPVPEEL 182  
QY 227 PGTFIGAV--SKLDYLKELGVNCTELMPCHEFNELEYSTSSKMFNGYSTINFFSPMTR 284  
DB 183 RGTAYAGLAHRAVVDHLISLGTAVELMRYHOFVHDGVLDDRLGLSNYWGYNITGFFAPHNG 242  
QY 285 YTSGINCGDAINEKTFEYREAHKRGIEVILDVFNHTAEGNENGPILSFKGVNDTTY 344  
DB 243 YALGTRG---QYSEEFKSMYKTLHEAGLEVLIDVYNNHTAEGNERGPTLSFGCIDNASY 299  
QY 345 YMLAPKG--EFTYNSGCGNTFNCNHPVYVROFIVDCLRYWTVMHVDGFRFDLASTIMTSGS 402  
DB 300 YRLVDGDMQHYYDTTGTGNSLLMRHPYLOLIMSLRYWTEMHVDGFRFDLASTLAROF 359  
QY 403 SLMDPVNYGAPLEGDMITTTGTPVTPPLMISNDPILGGVKTIAEAMDAG-GLYQVGO 461  
DB 360 HEVDRLSAF-----FDLIQDDPYISRKTLAEPMDVGEQGYQYGN 399  
QY 462 FPHMNVSEWNGKYRDIYRQFIKGTGFAEGFAECLGSPHLYQAGGRKPMHSINFCVAH 521  
DB 400 FP--QLMSEWNGKYRDAVDFWRAEDHSLGFEFASRLTGSSDLYQHSRRPRASVNFYTAH 457  
QY 522 DGFLLADLVYTNKKYKYNLPRGENNRDGENHNLNMGCEGEFARLSVKRLRKROMRNFVC 581  
DB 458 DGFLLRDLVSYNDKRNHNEARGEDNDGESHNSWNCGAEGGTKDPAYRELGRQRNPLAT 517

QY 582 LMSQGVPMFMYGDEYGHGTGKGNNTYCHDSYVYVFRWBDKKEQYSELARFCCIMTKFKPE 641  
DB 518 LLLSQGLPMICHGDELGTQGNNAAYCQDNEISWIDRLDGEORALLDFARRLALR-- 575  
QY 642 CEGIGLEDEPTAKRLQW-----HGHP-----GKPDMSNSRFY----- 675  
DB 576 -----ADHPLYLRRRRRFFHGETLTHADQPLPDLYWLLPDAREMTDDMORSDAHIVGVFL 629  
QY 676 --AFSMKDER-----QGEIYVAFNTSHLPVAVELPERA-GRWEPVVDGKRPADYDLT 726  
DB 630 NGDAIAEPDPRGRPVVDDSFLLNLSHWEPADEFRLPDAGYGERWTALVDTADP----- 682  
QY 727 DDLDRRA-----LTHQFSHFLXSNLYPMLSSVYTLVLRP 762  
DB 683 DGVDEAEHKAQTRLTYEQ-----RSLVLLSRP 710

Search completed: July 31, 2002, 12:14:40  
Job time: 389 sec

